

CAMBRIDGE  
APPROACHES  
TO LANGUAGE  
CONTACT

# Contact Languages

Ecology and Evolution in Asia

Umberto Ansaldo

---

CAMBRIDGE

This page intentionally left blank

## Contact Languages

Why do groups of speakers in certain times and places come up with new varieties of languages? What are the social settings that determine whether a mixed language, a pidgin, or a Creole will develop, and how can we understand the ways in which different languages contribute to the new grammar? Through the study of Malay contact varieties such as Baba and Bazaar Malay, Cocos Malay, and Sri Lanka Malay, as well as the Asian Portuguese vernacular of Macau, and China Coast Pidgin, the book explores the social and structural dynamics that underlie the fascinating phenomenon of the creation of new, or restructured, grammars. It emphasizes the importance and interplay of historical documentation, socio-cultural observation, and linguistic analysis in the study of contact languages, offering an evolutionary framework for the study of contact language formation – including pidgins and Creoles – in which historical, socio-cultural, and typological observations come together.

UMBERTO ANSALDO is Associate Professor in Linguistics at the University of Hong Kong. He was formerly a senior researcher and lecturer with the Amsterdam Center for Language and Communication at the University of Amsterdam. He has also worked in Sweden and Singapore and conducted fieldwork in China, the Cocos (Keeling) Islands, Christmas Island, and Sri Lanka. He is the co-editor of the Creole Language Library Series and has co-edited various journals and books including *Deconstructing Creole* (2007).

## *Cambridge Approaches to Language Contact*

### *General Editor*

Salikoko S. Mufwene, *University of Chicago*

### *Editorial Board*

Robert Chaudenson, *Université d'Aix-en-Provence*

Braj Kachru, *University of Illinois at Urbana*

Raj Mesthrie, *University of Cape Town*

Lesley Milroy, *University of Michigan*

Shana Poplack, *University of Ottawa*

Michael Silverstein, *University of Chicago*

Cambridge Approaches to Language Contact is an interdisciplinary series bringing together work on language contact from a diverse range of research areas. The series focuses on key topics in the study of contact between languages or dialects, including the development of pidgins and creoles, language evolution and change, world Englishes, code-switching and code-mixing, bilingualism and second language acquisition, borrowing, interference, and convergence phenomena.

### *Published titles*

Salikoko Mufwene, *The Ecology of Language Evolution*

Michael Clyne, *Dynamics of Language Contact*

Bernd Heine and Tania Kuteva, *Language Contact and Grammatical Change*

Edgar W. Schneider, *Postcolonial English*

Virginia Yip and Stephen Matthews, *The Bilingual Child*

Bernd Heine and Derek Nurse (eds.), *A Linguistic Geography of Africa*

J. Clancy Clements, *The Linguistic Legacy of Spanish and Portuguese*

Umberto Ansaldi, *Contact Languages*

### *Further titles planned for the series*

Guy Bailey and Patricia Cukor-Avila, *African-American English*

Maarten Mous, *Controlling Language*

Jan Blommaert, *The Sociolinguistics of Globalization*

Bridget Drinka, *Language Contact in Diachronic Perspective*

# Contact Languages

## *Ecology and Evolution in Asia*

---

Umberto Ansaldo

*The University of Hong Kong*



**CAMBRIDGE**  
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS

Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,  
São Paulo, Delhi, Dubai, Tokyo

Cambridge University Press

The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

[www.cambridge.org](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9780521863971](http://www.cambridge.org/9780521863971)

© Umberto Ansaldo 2009

This publication is in copyright. Subject to statutory exception and to the provision of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published in print format 2009

ISBN-13 978-0-511-64163-3 eBook (NetLibrary)

ISBN-13 978-0-521-86397-1 Hardback

ISBN-13 978-0-521-68253-4

Cambridge University Press has no responsibility for the persistence or accuracy of urls for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

A i miei genitori,  
zia Assunta,  
ed alla memoria di zio Riccardo





# Contents

---

|                                                                 |                |
|-----------------------------------------------------------------|----------------|
| <i>List of maps</i>                                             | <i>page</i> ix |
| <i>List of tables and figures</i>                               | x              |
| <i>Series editor's foreword</i>                                 | xi             |
| <i>Acknowledgements</i>                                         | xiii           |
| <i>List of abbreviations</i>                                    | xvi            |
| <br>                                                            |                |
| 1 Introduction                                                  | 1              |
| 1.1 Introduction to contact language formation (CLF)            | 1              |
| 1.2 Research questions                                          | 3              |
| 1.3 The role of ecology in Asian contexts                       | 5              |
| 1.4 Theory of language and CLF                                  | 9              |
| 1.5 CLF beyond exceptional evolution                            | 13             |
| 1.6 Outlook                                                     | 16             |
| <br>                                                            |                |
| 2 The ecology of Monsoon Asia                                   | 18             |
| 2.1 Monsoon Asia                                                | 19             |
| 2.2 Sino–Javanese trade                                         | 22             |
| 2.3 The city-ports                                              | 29             |
| 2.4 Manpower in early Southeast Asia                            | 35             |
| 2.5 The Western impact                                          | 42             |
| <br>                                                            |                |
| 3 Linguistic ecologies of Southeast Asia                        | 52             |
| 3.1 Southeast Asia and the role of Malay                        | 53             |
| 3.2 Malay contact varieties                                     | 60             |
| 3.3 Introducing contexts of formation                           | 65             |
| 3.4 The role of Portuguese in Southeast Asia and Southern China | 73             |
| 3.5 Summary                                                     | 78             |
| <br>                                                            |                |
| 4 Methodological issues in the study of contact languages       | 81             |
| 4.1 The ideology of theory                                      | 83             |
| 4.2 Multilingualism and transmission                            | 87             |
| 4.3 Conclusions                                                 | 97             |

|     |                                                           |     |
|-----|-----------------------------------------------------------|-----|
| 5   | Contact language formation in evolutionary theory         | 99  |
| 5.1 | Competence, performance, and socialization                | 101 |
| 5.2 | Language evolution and contact languages                  | 105 |
| 5.3 | Functionalist assets for contact linguistics              | 114 |
| 5.4 | Conclusions                                               | 118 |
| 6   | Congruence and frequency in Sri Lanka Malay               | 122 |
| 6.1 | The SLM community                                         | 123 |
| 6.2 | Selection and replication in SLM                          | 128 |
| 6.3 | Freeing SLM from the chains of exceptionalism             | 137 |
| 6.4 | Final remarks                                             | 144 |
| 7   | Identity alignment in Malay and Asian-Portuguese Diaspora | 148 |
| 7.1 | The ecology of identity alignment                         | 149 |
| 7.2 | Multiple alignments in contact settings                   | 150 |
| 7.3 | Identity alignment and admixture                          | 160 |
| 7.4 | Conclusions                                               | 180 |
| 8   | Pidgin ecologies of the China coast                       | 184 |
| 8.1 | Socio-historical background of Europe–China relations     | 186 |
| 8.2 | The ecology of Macau and the Pearl River Delta            | 188 |
| 8.3 | Grammatical features of China Coast Pidgin                | 197 |
| 8.4 | The missing Makista link?                                 | 205 |
| 8.5 | Discussion                                                | 208 |
| 9   | Implications, conclusions, and new horizons               | 213 |
| 9.1 | Theoretical and methodological implications               | 214 |
| 9.2 | Conclusions and new horizons                              | 222 |
|     | <i>References</i>                                         | 229 |
|     | <i>Index</i>                                              | 253 |

# Maps

---

|                                                                                             |                |
|---------------------------------------------------------------------------------------------|----------------|
| 2.1 Monsoon Asia (with wind patterns)                                                       | <i>page</i> 20 |
| 2.2 Mongol Empire <i>ca.</i> AD 1300                                                        | 24             |
| 2.3 China: coastal regions and languages                                                    | 25             |
| 3.1 Contact Malay varieties and their main locations in the<br>Malay-Indonesian archipelago | 58             |
| 3.2 Cocos (Keeling) Islands                                                                 | 67             |
| 6.1 Malay communities in Sri Lanka                                                          | 126            |
| 8.1 Pearl River Delta                                                                       | 185            |

## Tables and figures

---

### Tables

|                                                                         |                |
|-------------------------------------------------------------------------|----------------|
| 3.1 Linguistic ages in Singapore (adapted from Lim 2007)                | <i>page 70</i> |
| 6.1 Sri Lanka Malay communities – sociolinguistic profile               | 127            |
| 6.2 Etymology of case markers in SLM                                    | 129            |
| 6.3 Functions of case in adstrates Sinhala and Tamil                    | 130            |
| 6.4 Thematic roles in SLM and adstrates Sinhala and Tamil               | 130            |
| 6.5 Assignment of Dative in SLM and adstrates<br>Sinhala and Tamil      | 131            |
| 6.6 Assignment of Accusative in SLM and adstrates<br>Sinhala and Tamil  | 132            |
| 6.7 Morpheme sources for TA markers in SLM                              | 135            |
| 6.8 Aspect in PDM varieties                                             | 136            |
| 7.1 Kinship terms in BM, Hokkien, and Malay                             | 165            |
| 7.2 Cocos Malay terms of address                                        | 165            |
| 7.3 Pronominal systems of BM, Bazaar Malay,<br>Cocos Malay, and Hokkien | 166            |
| 7.4 Reduplication patterns in Singlish and adstrates                    | 168            |
| 7.5 Polyfunctional ‘have’ in BM/BZM, Malay, and Sinitic                 | 171            |
| 7.6 Morphological material in Cocos Malay                               | 178            |
| 8.1 Contacts in the Canton region                                       | 195            |

### Figures

|                                                                                   |     |
|-----------------------------------------------------------------------------------|-----|
| 5.1 Sources of learning in a homogeneous and normative<br>linguistic environment  | 100 |
| 5.2 Sources of learning in a heterogeneous naturalistic<br>linguistic environment | 101 |
| 5.3 Internal and external ecology                                                 | 112 |
| 6.1 The Control feature                                                           | 131 |
| 9.1 Multilingual and monolingual ecologies                                        | 226 |

## Series editor's foreword

---

The series *Cambridge Approaches to Language Contact* (CALC) was set up to publish outstanding monographs on language contact, especially by authors who approach their specific subject matter from a diachronic or developmental perspective. Our goal is to integrate the ever-growing scholarship on language diversification (including the development of Creoles, pidgins, and indigenized varieties of colonial European languages), bilingual language development, code-switching, and language endangerment. We hope to provide a select forum to scholars who contribute insightfully to understanding language evolution from an interdisciplinary perspective. We favour approaches that highlight the role of ecology and draw inspiration both from the authors' own fields of specialization and from related research areas in linguistics or other disciplines. Eclecticism is one of our mottoes, as we endeavour to comprehend the complexity of evolutionary processes associated with contact.

We are very proud to add to our list Umberto Ansaldo's *Contact Languages: Ecology and Evolution in Asia*. This is, to my knowledge, the most comprehensive account to date of 'contact language formation' in Monsoon Asia, including Sri Lanka, the Malaysian-Indonesian region, and Southern China. Providing a rich socio-economic history of layers of colonization of various kinds in the region, by Asians and Europeans alike, the book sheds light on the spread of indigenous languages such as Malay and Chinese, which, incidentally, have not developed identical ethnographic functions in the new geographical spaces of their expansion. It articulates clearly the relevant ecological factors that favoured both the geographical expansion of some languages indigenous to Asia and the emergence of new varieties such as Bazaar, Baba, and Cocos Malay, especially the 'city-ports' in relation to which the actual agents of these particular evolutions are identified. It is likewise informative to find out where and how colonial European languages indigenized, producing 'pidgins', 'Creoles', or other less divergent local varieties, and even why the evolution of each language was not uniform from one colony to another.

The reader will certainly be curious to find out why the languages of the economically and/or politically more powerful did not always prevail. Some of the explanations are to be sought in the ways in which the European colonization

of Monsoon Asia not only varied internally, such as between the Portuguese, the Dutch, and the British, but also differed from the layers of the Asian colonization that preceded it. It is certainly as important to ask why the Asian language varieties that spread are the non-standard ones as to query why neither Arabic nor Hokkien or Cantonese became major lingua francas, when both the Arabs and Southern Chinese had been trade superpowers in the region before the arrival of Europeans in the sixteenth century. The ecologies of the contact and competition between Chinese and Malay, both before and during the colonization of the region by the Europeans, are especially very informative.

No less significant is the question of how 'contact language formation' in Monsoon Asia differs from or resembles what we have learned to date about language contact in the Caribbean and Indian Ocean, regions which have informed our hypotheses the most about 'language change under contact conditions'. Umberto Ansaldo contributes critically to the old debate about universalist, substratist, and superstratist hypotheses, highlighting the merits of the ecological and uniformitarian approach to the subject matter. *Contact Languages* is impressive in its eclecticism too. Not only does it fruitfully bridge views by various creolists, historical linguists, and language typologists, but it also shows how various theories of evolution outside linguistics enable us to better understand both the mechanisms and the ecological factors that drive language speciation. The theoretical bridge between the linguistic effects of colonization in Monsoon Asia and the hitherto better-documented cases of the emergence of Creoles both around the Atlantic and in the Indian Ocean is as commendable in this case as the comprehensive coverage of the relevant literature in various complementary research areas, which is quite in the spirit of CALC. The reader will find in this book several useful references to language contact outside creolistics, to language acquisition, to linguistic anthropology, to the philosophy of science, to language evolution, and to colonial history, among others. Experts and non-experts alike will appreciate especially the wealth of information that the book provides about its subject matter.

SALIKOKO S. MUFWENE, *University of Chicago*

## Acknowledgements

---

The research and thoughts that led to this book started many years ago and I fear I may not be able to recall everyone who helped me get to this point.

I first of all wish to acknowledge the lecturers who inspired me most in my undergraduate years at the University of Venice, in particular: Magda Abbiati and Maurizio Scarpari for Chinese and historical linguistics respectively, and Massimo Raveri for Asian philosophy. Next I wish to thank PhD supervisor and friend Maria Koptjevskaja-Tamm for her generous teachings in linguistic typology, research methods, and dissertation writing, and her continuous support in the past years.

Much gratitude goes to a large number of individuals in two communities where I have conducted fieldwork: the Cocos Malays and the Sri Lanka Malays. In the former, I wish to thank the following for their time and hospitality: Nek Aman (Cree Haig), Nek Harun (Capstan Benjamin), Wak James (James Benjamin), the late Nek Keti lakilaki (Yapat Carey), Nek Keti perempuan (Ebetie Sebie), Nek Fatima (Pirus), Nek Ashari (Medous Bynie), Wak Pitah, Pak Adam (Rabuhu Anthoney), Pak Aini (Radal Feyrel), Mak Aini (Yean Yapat), Marias Blacket, Pak Shafi (Zaitol Wallie), Morin (Catalina Capstan), Matt Macrae, Mak Yazid (Nollia Zaibie), Mohd Noor Jada, Rosita Oemar, Aini Radal and Aindil Minkom, in particular Pak and Mak Aini and their family for their warm friendship. In relation to the history of the Cocos Islands, I also wish to thank Ron Grant, Pearl Tan, John Clunies-Ross, Jack Carlsen, John Hunt, and Kenneth McPherson for their input. In Sri Lanka, my most sincere gratitude goes to the following individuals for their interest, generous support, and warm hospitality and friendship: T. M. Muhaj Hamin, B. D. K. Saldin, T. K. Azoor, Salma and Upali Peiris and all their family, B. Jalaldeen, M. T. M. Rihan, B. J. Faizal Cassim; others whom I thank for their interest and time include: in Colombo, Stanley Oemar, K. Moulana, Fathima Mohamed, Dalrene Burah, Kartini Mohamed, Eileen Dane, Mohamed Ismath Suhood, Romola Rassool, Fathima Shahdia Jamaldeen, Talisha Rozianthi Jamldeen, M. Naveed Jamaldeen, Dane Danish Suhood; in Hambantota, the late Proctor T. S. Doole, M. Doole, the late T. A. Amath, T. M. Farook, Barmatha Cassim; in Kirinda, J. M. Rafaidin, the late A. S. Damani (Nenek), the late A. D. Prenantha, N. P.

Alawudeen, R. M. Nahar, S. M. Nawas, Siti Fatima, T. F. Salma. I am indebted to Thiru Kandiah for providing the initial point of contact with the Sri Lanka Malays.

The following colleagues and friends have offered comments to various parts of this book: I thank Christopher Stroud for his suggestions on sociolinguistic theory, Wolfram Hinzen for his view on linguistic theory and the philosophy behind it, Perry Johansson for his comments on the historical section, Hugo Cardoso for his feedback on Indo-Portuguese, and especially Stephen Matthews for comments on the data and many inspirational discussions on language contact and typology over the years.

Other colleagues have helped with materials, suggestions and advice. I want to thank Sander Adelaar and Jim Collins who have made their immense knowledge on Malay linguistics available to me by sharing their works and views. I thank Alan Baxter for generously providing me with materials on Makista, and Clancy Clements for making available some early drafts of his book on the linguistic legacy of Spanish and Portuguese. I also want to thank Geoff Smith and Stephen Matthews for sharing with me valuable materials and views on China Coast Pidgin. More thanks for their ideas, suggestions and materials also go to: Aone van Engelenhoven, John Holm, Brian Joseph, Yaron Matras, Miriam Meyerhoff, Pieter Muysken, Sebastian Nordhoff, Scott Paauw, Rachel Selbach, and Uri Tadmor.

There are colleagues in many places and universities around the world who were kind enough to listen to the ideas developed in this book at seminars, workshops, conference presentations, and invited talks over the past few years. I remember in particular the colleagues in Stockholm, Nijmegen, Leiden, Giessen, Edinburgh, Coimbra, Cape Town, Western Cape, Colombo, Kelaniya, Hong Kong, and Seoul, for interesting discussions.

At the University of Amsterdam, I would like to thank Kees Hengeveld for his constant support throughout the project, Enoch Aboh for many constructive discussions on Creole languages, as well as Norval Smith and Hans den Besten for being always available with expert advice.

I also acknowledge valuable assistants who helped with transcribing and organizing the Malay data and metadata: Diane Chang, Christine Angelique Kang, and Konrad Rybka.

For hospitality during my sabbatical year, which I devoted to this project, my gratitude goes to my good friends Charles Goddard, Tom Mitchell, and Martin Webster.

There are those I cannot mention: they know why.

The role of Salikoko Mufwene cannot be overestimated, having been most supportive of this project from the earliest phases. His numerous comments, suggestions, and ideas, and his constructive criticism on content, style, and



organization have been fundamental in arriving at what is hopefully a clear and complete product.

I also wish to thank Helen Barton and the whole production team at CUP for their efficiency.

My deepest and most emphatic gratitude must, however, go to the one person in particular who, more than anyone else, has supported this book project in too many ways to mention. Lisa Lim has contributed phonetic and sociolinguistic expertise, as well as professional advice as linguistic and copy editor, and has been a valuable fieldwork and life partner throughout. She has also afforded me unique, insider's perspectives on the ecology and evolution of new languages derived from her Peranakan heritage and Singlish competence. 'Thank you', Lisa, is barely enough, I know.

The following grants should be acknowledged: 'Contact languages of Southeast Asia: The role of Malay' (R-103-000-020-112), Academic Research Grant, National University of Singapore, for fieldwork on Cocos Malay and (partially) Sri Lanka Malay; 'The documentation of Sri Lanka Malay: Linguistic and cultural creolization endangered', Volkswagen Stiftung initiative for the Documentation of Endangered Languages (DoBeS), for fieldwork on Sri Lanka Malay; and, last but not least, 'Contact languages of Asia', The Netherlands Organisation for Scientific Research (NWO), for precious time to write this book.

*Amsterdam, November 2008*

Much of the content of this book has not appeared in press before but I would like to acknowledge the following: parts of chapter 4 have appeared in Ansaldo and Matthews (2007); a version of chapter 6 was the basis for Ansaldo (2008); for some sections of chapter 7, I am indebted to earlier work appearing in Ansaldo and Matthews (2004), Lim and Ansaldo (2007), and Ansaldo, Lim and Mufwene (2007); an earlier version of chapter 8 constitutes the basis of and is informed by later developments of Ansaldo, Matthews and Smith (to appear). I thank all these co-authors for the inspiration they have provided.

## Abbreviations

---

|      |                            |
|------|----------------------------|
| ABL  | Ablative                   |
| ACC  | Accusative                 |
| ANIM | animacy                    |
| BM   | Baba Malay                 |
| BZM  | Bazaar Malay               |
| CCP  | China Coast Pidgin         |
| CL   | classifier                 |
| CLF  | contact language formation |
| CM   | Cocos Malay                |
| CONT | continuous                 |
| DAT  | Dative                     |
| DEM  | demonstrative              |
| DEF  | definiteness               |
| DOM  | definite object marker     |
| FT   | foreigner talk             |
| LC   | linguistic citizenship     |
| LOC  | Locative                   |
| MM   | <i>Macau Monograph</i>     |
| N    | noun                       |
| NEG  | negation                   |
| NP   | noun phrase                |
| NOM  | nominative                 |
| NUM  | number                     |
| O    | object                     |
| PASS | passive                    |
| p.c. | personal communication     |
| PDM  | Pidgin-derived Malay       |
| PFV  | perfective                 |
| PK   | Papia Kristang             |
| PL   | plural                     |
| POSS | Possessive                 |
| PP   | prepositional phrase       |

|         |                           |
|---------|---------------------------|
| PROG    | progressive               |
| PST     | past                      |
| Psych V | psychological verb        |
| RG      | <i>Redhaired Glossary</i> |
| SVC     | serial verb constructions |
| SLM     | Sri Lanka Malay           |
| S       | subject                   |
| TA      | tense and aspect          |
| TM      | typological matrix        |
| TMA     | tense, mood, and aspect   |
| TRANS   | transitive                |
| V       | verb                      |
| VP      | verb phrase               |



# 1 Introduction

---

You are sitting in the tropical garden of a café nursing your thirst with a cold beer when you hear the following conversation:

- *Eh41 yu got watch PCK las nait or not?*

- *Ya lor55. Wa lau he kena sen English class! His English so lau ya one meh55?*

- *Ya wat21! Gamen say mas spik gud English ma21. Don play-play! Oderwise people ting yu dam chingchong la21.*<sup>1</sup>

You understand some of the words, but not all of them, because, although they sound English, some are pronounced in ‘strange’ ways. More importantly, you cannot quite make sense of all that is being said. So was it really English you heard or did you just imagine it? If it was, what was going on with the grammar? Because, surely, there was something wrong somewhere. And what about those little exclamations at the end of each sentence with such strange intonation? You might think that you have just heard some rather ungrammatical English being spoken. Or was it Chinese with some English words in it? According to some linguists, it was neither. What you have just heard is a language variety known as Singlish, or Colloquial Singapore English in academic circles. Singlish is a native variety of Singapore whose lexicon and grammar in part derive from English, Chinese and Malay. It is spoken by a majority of the young generation of Singaporeans, who relish it as their native tongue. It is officially labelled not ‘good English’ by the government and active steps have been taken in order to discourage its use. Educators and politicians in Singapore speak negatively of it, yet use it in the classroom and in political rallies. Ask five different language experts (or linguists) what they make of such a language and you are likely to get four different answers. This is the world of contact language formation, i.e. the evolution of new languages in multilingual environments. This book deals with the many issues – social, historical, and grammatical – that such languages raise.

## 1.1 Introduction to contact language formation (CLF)

Why do languages change over time? What leads a community of speakers to adopt a new language or, even more interestingly, to develop a new one?

Language change – in particular contact-induced change – has developed into one of the most fascinating areas of linguistics since Weinreich's (1953) seminal work. There are several reasons to justify the appeal of the field: perhaps the most striking is its multidimensionality, spanning as it does traditional sub-disciplines of linguistics, namely sociolinguistics, historical linguistics, structural analysis, and the ideology of language. In addition, understanding CLF, the topic of this book, holds the key to some of linguistics' most fascinating questions: How does a new grammar (and a new language) emerge? What are the socio-historical conditions that trigger it and what factors determine the grammatical outcome? What is it that speakers do in situations of language contact and why do they do it? Are there constraints on the creativity of language or is CLF completely unpredictable?

In many parts of the world, children and adults function in multilingual environments where several languages are spoken, perhaps in different domains, such as one language for the home (used predominantly in the family or one's ethnic group), one or more language(s) for the market place, and yet another language for official interethnic communication. Often these languages are mastered very early on and learnt in an informal environment, that is, without the support of school or other institutions. The strict sense of grammatical norm that speakers of Western cultures have become accustomed to in the past few centuries is not felt to the same degree in these communities, where variation between speakers along social class, ethnolinguistic group, religious sphere or generation may be very pronounced. In these multilingual environments switching between and mixing of different languages is the norm. Even in Western societies, mixing of codes can be readily observed in bilingual environments, such as within immigrant communities (think of *Gastarbeiterdeutsch* 'foreign workers' German', the variety spoken by predominantly Turkish immigrants) or in border areas (e.g. Italy's Northwest, where the Italian-German variety known as *Alto-Adige* or *Südtirol* is spoken).

In certain times and places, practices of code-mixing have eventually led to the formation of a new language. Perhaps the most familiar evidence comes from Creole languages, typical products of the displaced communities that emerged out of European colonial exploitation of, in particular, West African, Caribbean, and Pacific regions. But population movements, forced and free migrations, and population admixtures have in fact recurred in human history long before the recent European colonial era (fifteenth to twentieth centuries) and are still ubiquitous today. Indeed we can detect them in all regions of the globe for which we have good historical documentation.

The first time I experienced the fascination of a mixed language and culture was in Melaka, Malaysia. As a second-year student of Mandarin Chinese, I had already experienced the marvel of 'variation' when I encountered Cantonese in Hong Kong and later Minnan (or Hokkien) in Penang. I had expected to find

other Sinitic languages to be somewhat similar to the Mandarin I had learnt but I was instead struck by completely different grammatical systems; the Sinitic varieties were not mutually intelligible, which rendered the application of the label ‘dialects’ completely useless in their case. In Melaka I discovered another dimension of marvel in the Peranakan<sup>2</sup> culture, with its intriguing fusion of Hokkien and Malay elements, in food, rituals, language, and other cultural domains. When and how had this happened? Why had the Chinese *Babas* and the Malay-Indonesian *Nyonyas* developed this new language known as Baba Malay?<sup>3</sup> And how exactly could one explain the grammatical outcome? I carried these questions with me during my graduate years in Stockholm and Hong Kong, and found other, fascinating phenomena, such as Makista (or Macanese), the variety of Portuguese infused with Sinitic elements in Macau, which was, until recently, a well-preserved example of a city-port in which East and West met. The encounters that produced these mixed language varieties also produced hybrid architectural styles and innovative cultures that offer a fascinating insight into cultural and linguistic admixture.

My first academic appointment was at the National University of Singapore, where I had the opportunity to study in depth the incredible multilingual ecology of Singapore, in which Singlish and Bazaar Malay have emerged. Also, thanks to a university research grant, I was able to conduct fieldwork on Cocos Malay and Sri Lanka Malay, two interesting contact Malay varieties of the Indian Ocean. In more recent years I have been working with the Sri Lanka Malay speech community as part of Volkswagen Stiftung’s initiative for the documentation of endangered languages (DoBeS). The present book reflects the knowledge that I have gained to date from this rich linguistic experience, knowledge based on case studies that I conducted on salient aspects of CLF in these Asian contexts. Some of the restructuring is still in process and was investigated first-hand in fieldwork I conducted over the years. Other varieties are no longer spoken; thus the contact scenarios had to be reconstructed through historical and typological analysis. I discuss all of them in the following pages.

## 1.2 Research questions

This book aims at providing a conceptual framework, as well as a number of case studies, for exploring one underlying theme: the interaction between social factors and grammatical features in the process of CLF. In order to further our understanding of how new languages emerge, it is important to study the division of labour between ‘external’ and ‘internal’ factors in the evolution of new grammars (Mufwene 2001). This means that we need to bring together observations of a socio-historical nature with grammatical (here: functional-typological) analysis, laying emphasis on the former as the main force behind

CLF, as advocated by Thomason and Kaufman (1988). Specifically, throughout this book, I attempt to offer some answers to the following questions:

1. What kinds of ecologies trigger the emergence of new languages?
2. What is the relationship between input languages and the new grammar?
3. How do speakers use languages in multilingual environments?
4. Why do new languages evolve?
5. Which are the theories of language that can be reconciled with the lessons learned from CLF?

Following in the footsteps of Thomason and Kaufman (1988), I seriously consider the claim that languages change because speakers change them, by looking at both speakers' agency and grammatical systems in accounting for CLF in Asian contexts. Through the study of cultural and linguistic contact in Asian contexts, I further explore the relationship between the formation of new identities and the emergence of new languages, already established by Le Page and Tabouret-Keller (1985) in the case of the West Indies. Moreover, the case studies in this book show that a unified approach to the processes of language change and language maintenance clarifies and improves our current understanding of language change, in line with an evolutionary approach to language change as exposed in Croft (2000) and Mufwene (2001). In evolutionary terms, language change and the lack of it are seen as processes that can be explained by the same cognitive mechanisms of selection and replication of utterances. Finally, this book stresses the importance of ecology, in the sense of the environment in which CLF takes place ('external ecology'), as well as the systemic interaction of the linguistic codes in contact ('internal ecology'), as already pioneered by Mufwene (2001, 2008). While at the University of Amsterdam, I have enjoyed the benefit of exchanging ideas with colleagues whose expertise lies in language contact in different regions of the world, in particular West Africa, the Caribbean, and South Africa. It is in the hard work of these (and other) colleagues before me that I find inspiration to present what I hope will be novel, useful materials to advance our understanding of the creative linguistic behaviour typical of usually multilingual diasporas around the globe.

In the approach taken in this book (see chapters 4 and 5), the underlying assumption is that CLF, including the development of pidgins and Creoles, or mixed languages in general, is a natural outcome of multilingual environments in which languages are learnt in informal settings (the home, the work place, the market), as opposed to the school. This is particularly so in the case of displaced communities – diasporas that can be regarded as 'minorities' or peripheral within the geopolitical contexts in which they are situated. I will show that in polylectal speech communities, external factors such as (a) size of the community, (b) type of intra- and intersocietal networks, and (c) multilingual practices actively influence and, for the most part, determine the outcomes of



contact situations. ‘Speech community’ is intended here in the sense of a web of social networks (Milroy 1987; Milroy and Milroy 1992) with group-internal relationships as well as intercommunal relationships that, to various degrees, determine the types of linguistic attitudes that define the community.

Cutting across these parameters is the typological dimension, i.e. the similarity or difference between the grammatical systems in contact (‘congruence’ in Mufwene 2001). This dimension is of the utmost importance as I assume that it is only by relating grammatical analysis to points (a) to (c) above that a complete picture of CLF can be achieved. In this sense, I aim to explain patterns of CLF as essentially not divorced from observations derived from historical linguistics, language variation and change, areal typology, and sociolinguistics (e.g. Siegel 2004; Heine and Kuteva 2005; Ansaldo, Matthews and Lim 2007).

The case studies presented in this book offer substantial evidence for the fact that CLF is the natural – though not necessary – outcome of population admixture in shifting socio-historical contexts. The findings present compelling evidence in support of the following claims:

- language use is inherently creative;
- multilingualism and/or casual transmission support innovative use of language;
- the typological kinship of the input languages determines to a high degree the grammatical outcome;
- CLF goes in parallel with new identity formation;
- CLF is not the consequence of a negative process (i.e. failure to acquire a target language) but rather of a constructive process involving speakers’ agency; and
- CLF happens most often in minority groups, displaced groups, or otherwise geopolitically peripheral communities, who are not necessarily disempowered but often in a position of socio-cultural brokers.

### 1.3 The role of ecology in Asian contexts

If we acknowledge the role that ecology plays in the evolution of new varieties, it is extremely rewarding to step away from the traditional contexts in which the dynamics of CLF have been heavily researched so far. It is obvious that the lessons learned from these encounters, no matter how valuable, have limitations. In terms of both external and internal ecology, we have been looking at encounters between Western European languages and West African or Pacific languages in colonial environments within which the power relation is in favour of the Western party. This means that, in terms of linguistic typology as well as historical and social dynamics, we still have not yet explored the full range of possibilities in which CLF can occur.

In this book then, I discuss CLF in East, South and Southeast Asia, focusing on the role of Malay and Chinese in various encounters before, during, and after the European colonial era (fourteenth to twentieth centuries AD). The geographical area I cover extends from the South China Sea to the eastern half of the Indian Ocean, including the Indonesian archipelago. As has often been pointed out in traditional historical literature (e.g. Reid 2000), this region is first and foremost a climatic zone, ruled by the two monsoon seasons that determine the lifecycle of its populations in all significant aspects of life; it is also known as Monsoon Asia. Malay and Chinese varieties played an important role throughout the history of the region as the languages associated with the economically most powerful and influential territories, a status made more evident by the spice trade between the Maluku Islands and China (see [chapters 3 and 4](#)). This offers an opportunity to observe internal and external ecologies that are different from the better-known settings in which CLF has been studied, i.e. the Atlantic and Pacific regions.

Take the case of China Coast Pidgin (CCP), for example, the oldest attested Chinese–English pidgin we know of: here the linguistic encounter involves Cantonese and English. The former is a strongly isolating language; that is, it makes no use of inflections for person, number, and tense. Also importantly, its sound system is radically different from that of English, especially in its use of lexical tone. Moreover, the Chinese were by no means colonial subjects in the formative period of CCP. Quite the contrary, as I will argue, they were in a position of power, determining the local conditions of trade between them and the English. What are the social dynamics of such encounters? Was CCP the result of the Chinese failing to master English or was it the outcome of attempts by Western merchants to master Chinese? Where does the agency lie in the formation of this new language and how do we make sense of its grammar?

Ecological diversity is also conspicuous in the evolution of contact Malay varieties. For example, there were contexts of relative isolation, in which varieties such as Cocos Malay emerged that were influenced by pre-existing lingua francas, just as there were settings of intimate contact, where varieties evolved that were influenced by adstrate languages, as in the case of Baba Malay, which bears Chinese influence. The consequences of such ecological variation lie in the different grammatical outputs of the language restructuring process. Moreover, there are cases where contact spanned a very long period of time, such as that of Bazaar Malay, which was attested for roughly a millennium, and these offer us a new perspective on the nature of pidgins and lingua francas. Finally, the patterns of multilingualism observed in certain Asian ecologies are quite unlike what we know of, say, the Caribbean region, again providing us with valuable comparative material. This is, for example, the case of Sri Lanka Malay, a language formed through the practice of widespread trilingualism (see [chapter 6](#)).

As suggested above, using the case studies presented in the book, it will be possible to tackle the following questions in relation to CLF in Monsoon Asia:

- Does contact language formation in the Asian context present us with novel evolutionary patterns?
- What are the dominant external factors in Asian ecologies?

In order to answer these questions properly, it is important to appreciate fully the nature of the ecological factors that define the various contexts in Monsoon Asia, which are the following:

1. Slavery and manpower: the study of slavery and manpower in different regions and historical phases helps us understand the patterns of indentured labour and population movements that led to heteroglossic networks as seen in particular in the city-ports. This relates to questions of (i) mode of transmission, i.e. how new features and varieties got integrated with the already existing ones, and (ii) notions of ‘target language’ and feature pools available to speakers.
2. Economy and politics: seasonal trade routes determined by the monsoons gave rise to a number of multicultural and multilingual enclaves. How were the numerous languages negotiated and what were the outcomes of such negotiations in the different ports? And how did the transition of power from local networks through ‘interactive’ colonization (Faraclas, Walicek *et al.* 2007), involving trade, intermarriage and other forms of mutual exchanges, to exploitation colonization impact language change in the region?
3. Social relations: mixed cultures that arose in precolonial times from quasi-equal social relationships between the parties (e.g. marriages between Chinese merchants and local women in Malaysia, or communities that emerged from the mixing of Portuguese, Malays, and Chinese) provided ideal settings for the evolution of new languages.
4. Language varieties and linguistic codes: the typology of the languages involved, Malay and Sinitic, can show remarkable congruence at a superficial level. It is crucial to distinguish between written and spoken registers in both cases because there is a great divergence between the registers, and it was predominantly features from colloquial varieties that were involved in interethnic communication. It is also important to look at the appropriate varieties in each case, as there is significant regional variation. When looking at Malay influence, for example, we have to be aware of the variety of languages spoken in the Indonesian world and their significance in the past. When looking at Sinitic influences, we usually have to focus on southern varieties, Hokkien (Min) and Cantonese, typologically very divergent from Mandarin (or Standard Chinese).

In this context, I find it particularly enlightening to look at cases of CLF of precolonial times as a direct product of the social, cultural and political dynamics of Monsoon Asia. The development of pidgin Malay varieties such as Bazaar Malay (Adelaar and Prentice 1996) is a clear indication of the ubiquity of multilingualism and language admixture, and shows us that CLF can occur under historical circumstances that are not particularly extraordinary. Nonetheless, the question of why humans develop new languages over time, rather than speak like their neighbours or their ancestors, is a very intriguing one. As we will see, when social and historical conditions are ripe for a new community to emerge from the blending of previously existing ones, CLF is very likely to occur. The external conditions for CLF typically involve one or several of the following situations (see [chapters 2 and 3](#)):

- (a) trade between different communities, in particular where there is a stable space where the groups converge in order to trade, i.e. a market, a port, etc.;
- (b) interethnic marriage, especially when that happens in a neutral or ‘third space’ as in the case of the Hokkien Babas and Malay/Indonesian Nyonyas in Melaka (see [chapters 3 and 7](#));
- (c) resettlement as a consequence of migration, whether forced or spontaneous (indenture, economic, political, etc.), as observed in the Malay Diaspora of Sri Lanka (see [chapter 6](#)).

In newly formed ecologies created by such conditions, a community that has achieved a critical demographic mass and stability in the new space will often start identifying itself as a discrete cultural unit. As part of this process, there can be a societal convergence (both conscious and unconscious) towards a shared set of communicative practices that results in what appears as the evolution of a new language. For example, the case of the Peranakans of Melaka shows how a new community that arose out of intermarriages between Chinese- and Malay-speaking people came to develop a new, hybrid variety of Chinese–Malay, although both Chinese and Malay were fully available and functional in the ecology in which the Peranakans lived. The new language, Baba Malay, was not a communicative necessity or a failure in mastering either Chinese or Malay but rather an expression of the new cultural identity of the Peranakans, neither Chinese nor Malay. This is an example of why, in the remainder of this book, CLF is treated as an aspect, as well as a result, of new identity formation ([chapter 7](#)). Situations where language contact was extensive, but no new identity formation occurred, typically did not result in CLF. Extensive contact without CLF can be observed in ‘linguistic areas’ where the effects of contact result in typological convergence, as discussed by Gumperz and Wilson (1971) for the Indian village of Kupwar, and Ross (1996) for the case of New Guinea.

Another contribution of the case studies in this book is that, in all cases, there is substantial historical evidence to reconstruct a plausible ecology. In the cases of varieties still spoken today, such as Cocos Malay, Sri Lanka Malay and Bazaar Malay, I have been in the fortunate position of collecting first-hand data within the communities, not only relying on data elicited from informants but also recording natural language use and variation, as well as cultural data and oral history. In this sense, these case studies offer an invaluable opportunity to sketch theories based on sufficient information, rather than absence of data, as has often been the case for CLF in the Atlantic or the Mediterranean. For the varieties no longer spoken, there is often reliable socio-historical documentation for plausible reconstruction, as is the case for Baba Malay and Makista. This is discussed further in [section 1.5](#) as well as in [chapter 3](#).

## 1.4 Theory of language and CLF

In linguistics there are basically two views regarding the nature of language. In one view, our linguistic faculty is determined by innate structures that specify the range of things human language can do. In this sense, language is ‘generated’ by what has been identified as Universal Grammar, the biological endowment for language, or the language organ, which is common to all modern humans. In the other view, our linguistic abilities are determined by the functions of language; in this sense, language is the product of usage and, if there are universal aspects of language, they are to be found in general cognitive structures. Besides the generative view, developed by Noam Chomsky and his colleagues, and functionalist views, advanced in the work of Talmy Givón, among others, we also find intermediate views in which both positions are reconciled, as most clearly advocated in the work of Fredrick Newmeyer (see Newmeyer 2005).

Many linguists inspired by the first approach like to talk of language as an abstract entity with identifiable structural organizations that cross-linguistically have identical components. For example, they talk about ‘mechanisms’ of language contact and of ‘acquisition’ and the systemic transfer of features from one idiolect into another. However, in recent years, our increased understanding of linguistic diversity has proven time and again that universal features of language are hard to pin down, and it has been suggested that all aspects of syntactic structure may be language specific (Croft 2001: 364). For example, even the broadest, apparently universal features of language appear weak: Subject-Verb-Object (SVO) as the ‘default’ parameter (Kayne 1994), for instance, is severely questioned in typological literature (Dryer 1992; Hawkins 1994; see also [chapter 6](#)), and the generally assumed ‘universal’ noun-verb distinctions in parts of speech (Croft 2001) are questioned in particular in Gil (2000, 2001). Even more problematic is the obvious inherent diachronic change in language

that seems like a logical contradiction to the idea of a static, hard-wired linguistic ‘organ’ (Tomasello 2003). In fact, philosophers and psychologists, as well as evolutionary biologists, have become increasingly disenchanted by generative ideas of acquisition, and the case for innate structure is at this point at best weak, while there is an increasing amount of knowledge that argues against language-specific cognitive patterns (Tomasello 2003; Lieberman 2006).<sup>4</sup> Moreover, from the point of view of contact linguistics, despite attempts to identify structural constraints on contact-induced change, so far no such constraints have emerged (Koptjevskaja-Tamm to appear).

My work is very much part of the second trend and therefore incorporates a usage-based view of language and transmission; in this sense the context in which language functions is a primary domain of inquiry. The determining role of ecology in the process of CLF has been demonstrated in the evolutionary approach to language change developed in Mufwene (2001), in terms of ‘internal’ ecology (or the system-internal relations that can be observed when we abstract language from its social context) and ‘external ecology’, i.e. the socio-historical domain. When discussing the currents that affect the internal ecology of CLF, I focus on functional-typological features and cognitive aspects of language. Having established the primary role of external ecology, I assume that if any constraints are to be uncovered, they will be found in the socio-historical domain.

#### 1.4.1 *Evolution of grammar*

Language evolution can be seen as the cumulation of innovations occurring in individual speech acts of individual speakers (Mufwene 2001), and languages can be seen as complex systems “in which the processes that occur in individual usage events ... with high levels of repetition, not only lead to the establishment of a system within the individual, but also lead to the creation of grammar, its change and its maintenance within a speech community” (Bybee 2006: 730).

In functionalist and cognitive theories of grammar (part of the second position presented above), linguists do not isolate the structure of language from language use, as grammar is intended as the cognitive organization of a speaker’s experience with language (e.g. Givón 1979a; Langacker 1987). In both cognitive and functionalist theories, grammar is seen as a set of cognitive representations that rely on general cognitive abilities of categorization, generalization, representation, etc., and are therefore not language specific. While this view has not been the dominant one in linguistic theory, in recent years a number of trends have been accumulating substantial evidence in favour of a usage-based theory of language, in particular in the field of grammaticalization studies, language acquisition, and functional-typological linguistics (Bybee

1998, 2007). As noted in Bybee (2007: 8), a very important discovery that these approaches share relates to how living languages create new grammar. Research in these areas over the past two decades shows that grammar emerges naturally out of pre-existing lexical material in language use, as is abundantly documented in the literature on grammaticalization. Grammar in this view is a constantly evolving system and “there is every reason to believe that all existing grammar came about in just the same way we observe in the documented cases at our disposal” (Bybee 1998: 250). Grammar therefore can be seen as an emergent system, complex and dynamic, similar to other complex systems observed in biology (Bybee and Hopper 2001).

#### 1.4.2 *An evolutionary take*

The philosophical ideas on which my theoretical framework rests are simple: they are based on an evolutionary view of language as a dynamic, interactional phenomenon. As far as I am concerned, the evolutionary theory of language change has been convincingly presented in Croft (2000), and the ecological approach presented in Mufwene (2001) is largely compatible with it and explores further the explanatory potential of competition and selection in language change. In both approaches, which I briefly revisit in [chapter 5](#), three common elements stand out: (1) grammars are inherently variable, i.e. even speakers of the same speech community speak differently: therefore any grammar entails variable ways of saying things (i.e. it is either multilectal or multilingual); (2) change happens through selection and non-identical replication of linguistic features, and these processes lead to feature recombinations which differ from one speaker to another, which can foster divergence in some ecologies; and (3) selection takes place in a feature pool to which all speakers of a community contribute (Mufwene 2001).

In an evolutionary approach, language change – and therefore CLF – is a result of interaction between speakers in specific ecologies. We therefore focus on the two dimensions already mentioned above:

- i. External ecology: social, political, cultural, and historical patterns that may influence language use;
- ii. Internal ecology: the pool of features available to speakers in the multilingual space.

Through the study of external and internal ecologies, we can determine the nature of the basic evolutionary patterns: these are selection and replication of linguistic features in the process of CLF, which are typically sensitive to factors such as prestige, focussing, systemic or statistical frequency, structural regularity, and semantic transparency, subject to the founder principle (Mufwene 1998).

Considering my emphasis on external ecology, it is natural that I investigate contact language formation at the population level (E-language) rather than at the individual level (I-language). As discussed in [chapters 4 and 5](#), I believe that there is a methodological advantage in doing so because, if we acknowledge the fundamental role that society plays in the evolution of new grammar, it is the communal, societal output of language contact that we need to concentrate on. Moreover, I believe that in the study of past contact scenarios, we engage in a reconstruction process that can render a more informative picture at the level of E-language, or, put differently, may never be able to deliver a reliable reconstruction of an idiolect. This does not imply a firm belief in the reality of E-language vs. I-language, as I am comfortable admitting that they both are, in different ways, idealized academic constructs for the purpose of investigation. The former, however, seems to me more readily – and realistically – accessible than the latter, which has been prone to highly idealized and subjective constructions (Sampson 1997). Though it might be desirable that scholars investigating these two dimensions independently achieve similar results (Aboh and Ansaldo 2007), this may not be possible across the board. That it is useful to distinguish these domains of analysis becomes particularly clear when assessing our current understanding of the rate of contact-induced change. In E-language, change means propagation, and is clearly an external matter, which relates to type and size of society and moves gradually ([chapters 6 and 8](#)). In I-language, in a parametric view of grammar, for example, change may appear abrupt if interpreted as a matter of parameter setting (Lightfoot 1999). As shown in Mufwene (2001), this is disputable: idiolects do not change abruptly, any more than language ‘acquisition’ proceeds abruptly, as not all members of a class are concurrently affected by change. If we maintain a dichotomy between E- and I-language, we may thus arrive at opposite conclusions about rate of change (but see discussion in [section 4.2.1 and chapter 5](#)); in this sense it is important to note that these conclusions are not in conflict, as they pertain to different domains of enquiry, namely the study of language as a social construct vs. the study of language as an innate system.

### 1.4.3 *Ecological factors and system-internal analysis*

If we want to abstract the grammatical domain from CLF, in order to focus on the structural recombinations that occur in the interactions of idiolectal systems, it is useful to take into account the following aspects of the internal ecology:

- i. Token-frequency effects: here we identify features that are prominent because they occur with very high frequency in discourse, i.e. words that are used more often or commonly.
- ii. Type-frequency effects: here we look at features that are more prominent due to typological congruence between similar languages involved in the contact setting.



Both kinds of features cannot be fully understood purely at the system-internal level, even if it is there that we can isolate and study them. For example, (i) is not only a consequence of semantic-pragmatic salience but also an effect of statistical majority in a multilingual feature pool: the variety spoken by a dominant majority will be more present in discourse than others, and therefore more frequent. As for (ii), ‘congruence’, as defined in Mufwene (2001), indicates the degree of matching between different grammatical features; congruent features in different languages will favour maintenance in CLF, while a mismatch may undermine the survival of certain features. Congruence relates to the typological make-up of the feature pool.

In (re)constructing the internal ecology of contact, we therefore consider external factors such as numbers of speakers, their prestige, and their influence on internal aspects such as degree of congruence between systems, regularity, etc. This is why CLF is largely unpredictable, as the same relationship between comparable external and internal ecologies seldom repeats itself. All these factors can be seen at play in the formation of Sri Lanka Malay (SLM): in this ecology, the dominant language is Sinhala, both demographically and in terms of prestige, followed by Tamil, then Malay. Moreover, Sinhala and Tamil are similar to one another in many aspects of grammar, i.e. they are congruent, and this enhances both token- and type-frequencies of their features. For example, while Malay follows the Subject-Verb-Object (SVO) word order, both Sinhala and Tamil prefer the Subject-Object-Verb order (SOV). When we look at SLM, we see that the basic word order is indeed SOV. However, we also find maintenance of old Malay features in some areas of the grammar, which derive from the conflicting patterns in the adstrates that compromise selection. While in the noun phrase we see mostly Lankan grammar, in the verb phrase of SLM we find a number of Malay-like and innovative features: this can be explained by the fact that Sinhala and Tamil are far more congruent in the nominal domain than in the verbal domain. Simply put, the fact that speakers of SLM prefer verb-final word order, follow Lankan rules in marking grammatical relations, and structure events according to a combination of Malay and Lankan strategies is predictable from a close scrutiny of the social and typological aspects of all the languages involved in the multilingual ecology of Sri Lanka. In this sense, SLM is a completely ‘normal’ outcome of the sociolinguistic patterns and typological features that characterize the history of the SLM community.

## 1.5 CLF beyond exceptional evolution

In recent approaches to CLF, or language creation (DeGraff 1999), the field has become aware of a problematic tendency identified as ‘(Creole) exceptionalism’, i.e. the tendency to offer exceptional explanations in accounting for the formation of new languages that do not fit, or at times contradict, the generalizations on language and change derived from mainstream linguistic theory

(DeGraff 2001, 2003, 2005; Ansaldo, Matthews and Lim 2007). The effects of exceptionalism are illustrated in the quote below:

Another well-known instance of ‘non-genetic’ languages which did not arise as a result of a *normal and spontaneous* course of language evolution and have more than one source or parent are trade language, pidgins and creoles. Given the ongoing controversy concerning their validity as a special language type and the highly *specific social circumstances* which warrant their emergence, we choose not to expand the present volume into this field. (Aikhenvald and Dixon 2006: 11, fn.7; my emphasis)

‘Exceptionalism’ is the belief that there is something not ‘normal’ or natural about the evolution or formation of a new language. Perhaps this can be explained by the fact that Western linguistic ideology has for the longest time promoted maintenance and purism as ‘normal’ aspects of phylogeny; or maybe there is even a trace of postcolonial intellectual guilt to be detected here, in the exceptional nature attributed to anything arising from the conceptual space of Western colonial history. Whatever it may be, quotes such as the above justify the necessity of a volume such as this one, which aims to integrate the lessons learnt from the study of CLF with general linguistic theories. As we will see, CLF is as normal and spontaneous as any other linguistic process. And the social settings in which it occurs are by no means rare and typify much of human history, including the present time of globalization. In this sense, maintenance and innovation are not seen as ‘normal’ versus ‘abnormal’ patterns in language evolution, but rather as opposite poles on a continuum characterized by the same dynamics of language use.

The exceptional nature ascribed to contact languages and Creoles, I argue, derives from a biased understanding of the motivations behind language change (but cf. Baker 1990; Arends, Muysken and Smith 1994). Tutored acquisition in monolingual settings is quite rare in the history of the world while multilingual practices and casual transmission of language have been more prevalent by far in many of the world’s societies, a recurrent theme in much of Mufwene’s uniformitarian approach (2001–2008). It has been proposed that exceptionalism is partly the result of a field that has so far relied heavily on studies based on situations dominated by Western colonial powers and the ecolinguistic environments they create (DeGraff 2003; Ansaldo and Matthews 2007). In that field, Eurocentric notions that originate in nineteenth-century nation-state ideology have shaped, and still influence, our views on linguistic classification and language purism. These notions, often supported in traditional philological and historical linguistic environments, fail to capture the nature of the relationships and interaction of different languages in other parts of the world, as already pointed out in a number of studies on languages in Australia (Dixon 1997) and Laos (Enfield 2003). The limitations implicit in such a field are at least two:

1. The input–output relation. Typically, we look at encounters between isolating languages (e.g. West African) and fusional ones (Standard European) where the latter occupy a position of power. With such constant typological and social aspects, we predict similar outputs (albeit with variation) to arise from the contact environment. As mentioned in [section 1.3](#), when we shift our enquiry to Asian ecologies, the social and typological patterns change. As we will see in [chapter 2](#), for example, the type of ‘slavery’ observed in Asian contexts differs from what we know of the Caribbean region, leading to different social dynamics. Likewise, the grammatical systems in contact in this region include Chinese, Austronesian and South Asian varieties, thus offering a new perspective on the effects of typological admixture. The formation of Sri Lanka Malay, for example, shows a case of typological shift from isolating to agglutinative morphology, something that has been considered rare in CLF so far, which has led a number of linguists to formulate hypotheses relating the evolution of morphology – regarded as ‘complex’ structure – to time-depth (see McWhorter 2005).
2. The idealized transfer model. From a Western perspective, ‘normal acquisition’ involves learning in a scholastic environment, with a clear notion of standard language, and, often, very little linguistic diversity. This view has deeply influenced current theories of language transmission. But, crucially, compulsory education for the masses, standardization and monolingualism are part of what is a recent, Western European reality. They are not typical of human history, nor common across the world. The study of past, and to some extent present, Asian ecologies clearly shows that frequent and widespread multilingualism (or multilectalism) is the norm, and that casual, rather than formal, transmission is very frequent. Observing such contexts can add valuable knowledge to our understanding of patterns of second (and third) language acquisition in diverse multilingual contexts (cf. Siegel 2004). Again, the case of Sri Lanka Malay shows how typological congruence in the multilingual pool can go a long way in explaining the types of admixture that speakers undergo as a consequence of being heavily exposed to languages such as Sinhala and Tamil during the formative years of the Malay Diaspora.

The general assumption in this study is that language contact, i.e. the process of mixing of originally separate linguistic systems to create a somewhat new one, is the norm rather than the exception in the evolution of language. Social and grammatical hybridity, which I treat as the result of new identity formation, is what defines the most fundamental aspects of the continuously changing, heteroglossic communities of our world (Whinnom 1971; Croft 2001; Ansaldo 2003a). This hybridity needs to be understood in terms of degree: depending on the social and the linguistic type of the groups involved, the

output of a contact situation may resemble or depart from the input varieties to varying degrees. This is clearly illustrated by comparing different scenarios of CLF, such as Sri Lanka Malay and Cocos Malay, both varieties developed in the Malay Diaspora. In the former, in an ecology within which Sinhala and Tamil – Sri Lanka’s dominant languages – are spoken, we see a process of CLF in which the socio-historical conditions trigger a restructuring process that leads to a language that looks more ‘Lankan’ than Malay. In the latter, an ecology within which various Malay-like languages are spoken, and where there is little socio-cultural pressure to trigger change, the grammatical output will be less divergent from the input (see [chapters 6 and 7](#)).

## 1.6 Outlook

This book is about the formation of new languages and new cultures and identities. Contact between humans from different ethnic and linguistic groups is one of the dominant factors for the spread of material and social cultures. Many modern cities of today rest on the remains of what used to be ancient markets, border crossings or common waterways, where practitioners of different cultures and languages came to trade. Many such spaces still exist in the non-Westernized world, and language contact as well as contact language formation are still very much part of our human landscape. It is through contact and fusion that the most creative aspects of humanity can be discerned, as the new linguistic repertoires that emerge from CLF clearly show; in this sense, contact-induced change is a very ‘normal’ practice for humans, and an ancient one. It is the kind of language maintenance which requires the compilation of vocabularies, elaboration of spelling and grammatical norms, schools and academic institutions and an army of fierce language purists that appears abnormal from a human communicative point of view. Therefore, in order to understand fully the nature of human language and communication, we need to first understand the natural, spontaneous and creative use found in multilingual, untutored environments, of which CLF is a central part.

## Notes

- 1 The conversation is represented here not through standard English spelling, which would create the illusion of complete intelligibility between the language in question and English, nor through IPA, which would require linguistic training in order to be properly appreciated. In this way I hope to allow the general reader to access this short text and get the ‘feel’ of being exposed to a typical ‘Asian English’ variety. The numbers indicate pitch levels, where 5 is a high tone, and 1 is a low tone. (See chapters 5 to 7 for notes of Singlish grammar.) The conversation (courtesy of Lisa Lim) goes something like this:

-Hey, did you watch *PCK* [*Phua Chu Kang*, a local sitcom] last night?

-Yes! Wow, he was sent to English class – is his English really so bad?

-Of course! The government says one must speak good English. Don't fool around. Otherwise people will think you are really Chinese-educated.

2 This is a term used to refer to the descendants of Chinese merchants and Malay or Indonesian women born in Southeast Asia.

3 *Baba* refers to the male Peranakans, *Nyonya* to the womenfolk.

4 The innate language faculty has been recently reduced to one feature: 'recursive merge' (Hauser, Chomsky and Fitch 2002). This is a well-established feature of complex systems not unique to language (Edelman 1992).

## 2 The ecology of Monsoon Asia

---

*Moors from Cairo, Mecca, Aden, Abyssinians, men of Kilwa, Malindi, Ormuz, Parsees, Rumes, Turks, Turkomans, Christian Armenians, Gujaratees, men of Chaul, Dabhol, Goa, of the kingdom of Deccan, Malabars, and Klings, merchants from Orissa, Ceylon, Bengal, Arakan, Pegu, Siamese, men of Kedah, Malays, men of Pahang, Patani, Cambodia, Champa, Cochin China, Chinese, Lequeos, men of Brunei, Luçoes, men of Tamjompura, Lave, Banka, Linga (they have a thousand other islands), Moluccas, Banda, Bima, Timor, Madura, Java, Sunda, Palembang, Jambi, Tongkal, Indragiri, Kappatta, Menangkabau, Siak, Arqua (Arcat?), Aru, Bata, coutry of Tomjano, Pase, Pedir, Maldives.*

On the people who come to trade in Melaka; Tomé Pires (1515: 268)

Asia is a region with a rich past and a very diverse range of cultures, a significant part of which is well documented due to the long history of civilizations that shaped the region. This is especially so for the history of China as well as South and Southeast Asia, documented since ancient times in local chronicles and imperial histories, as well as by Western observers. The region defined in [chapter 1](#) as ‘Monsoon Asia’ is particularly interesting to us, characterized as it is by a history of grandiose naval expeditions, such as the imperial voyages of the Chinese admiral Zheng He, who at the dawn of the fifteenth century had travelled as far as the eastern coast of Africa with a massive fleet. This region was the theatre of prolonged and intense patterns of maritime trade, patterns that had been going on for many centuries and were known in Europe in ancient times, most importantly revolving around the ‘Spice Islands’. Throughout its history, we can observe the rise (and fall) of vibrant, multicultural coastal communities such as Melaka,<sup>1</sup> some of which are, to this day, global centres of trade (for example, Singapore). It is in these patterns of discovery, commerce and urbanization that different peoples meet and shape a crucial part of the linguistic history of the region, a phenomenon existing already in precolonial times. It is therefore necessary to take a close look at the history and geography that define this region and set the stage for the human interaction within which new communities – and languages – arise.

This chapter will illustrate the vibrant cultural sphere of Monsoon Asia, defined by networks of trade connecting a number of larger empires and smaller

sultanates. Political relations such as the protectionism that China offered to a number of vassal states provided periodic exchanges of goods, information and technology, as well as constant exchanges of envoys and visiting parties. The spread of religious doctrines, in particular Islam and Buddhism, was supported by a network of travelling scholars as well as numerous religious academies throughout the region, where scholars gathered around the teaching and translation of texts. All this was favoured by the climatic pattern of the monsoons. Bond servants, a common cultural feature of the region, also led to cultural admixture and, later, to population movement, and constituted, together with the intense commerce in spices and other valuable goods, the dominant force behind the establishment of self-regulated, quasi-egalitarian, multicultural coastal communities, in particular the city-ports. As we will see, it is in the ecology of these city-ports, within which cultures blend and new identities are negotiated, that new languages evolve over time.

## 2.1 Monsoon Asia<sup>2</sup>

The geographical region that this book covers in terms of contact language formation is in fact not ‘Asia’ at large, far too wide and rich in history and traditions to be treated in a single volume. The communities and the languages described here are situated in South China, peninsular Southeast Asia, and the eastern half of the Indian Ocean, a region that, as noted in [chapter 1](#), is sometimes referred to as ‘Monsoon Asia’ (see [map 2.1](#)). Since oceans and seas have often been recognized as fundamental structures for the understanding of human history, in this chapter I present the history of the region as defined by its waters (Pearson 2003). I focus especially on the structure and the dynamics of the Southeast Asian peninsular region, as this is, in many ways, the heart of Monsoon Asia.

The geographical area discussed in this volume can be visualized as extending from the Indian subcontinent in the west all the way to the southeastern corner of China to the east. Its northern and southern boundaries are marked by the tropic of Cancer and the tropic of Capricorn respectively. In viewing a region, it is its geophysical properties that define a number of constraints on patterns of human movements that are relevant for our understanding of history. In the case of bodies of water, special attention is given primarily to winds, as these determine the range of movements possible, as well as to choke points, such as straits, channels, and islands, as possible stops on the journey between places (Pearson 2003). In this respect, the most defining structural aspect of the Asian region discussed here is clearly the pattern of monsoon winds. This is because, unlike the Atlantic, for instance, which is essentially characterized by one and the same regular wind pattern throughout the year (known as the ‘trade wind’), the Indian Ocean is controlled by two regular wind patterns. These

are the monsoon winds (from Malay *musim*, ‘season’), which blow from the southwest from May to September, and from the northeast from November to March (map 2.1). Starting in the Arabian Gulf as early as April, the southwest monsoon moves across the Indian Ocean and then into the South China Sea, peaking in strength around July and then starting to weaken throughout August and September. By October, the northeast monsoon starts blowing from the South China Sea in the opposite direction, engulfing basically the whole of Southeast Asia by the month of January. It eventually weakens until the system starts again in April with the southwest monsoon.



Map 2.1 Monsoon Asia (with wind patterns)



The influence of the monsoon patterns on trade in the history of the region is of utmost significance. Throughout the monsoon region, peoples at sea were always assured of availability of regular routes to distant markets as well as a return home, no matter how far they travelled, a fact that encouraged long-range voyaging. This regular pattern of winds has supported and encouraged economic, cultural, and diplomatic exchanges in the Indian Ocean since the dawn of history, bringing into contact the South China Sea with the Indonesian archipelago, the Indian subcontinent, the Arabian Peninsula, and the eastern and southern African coast (McPherson 1993: 124).

The monsoon pattern also encouraged the establishment of temporary settlements, the result of traders needing to wait for the change in monsoon in order to return home. One of the consequences of this natural rhythm was the establishment of relatively stable overseas communities, as can be seen in the case of the city-ports of Melaka, Batavia (Jakarta), and Macau, among others. While new ports and new communities were being created, ties with the country of origin were not necessarily cut, as there was ultimately, in most cases, an intention to return home. Therefore, cultural ties between new and original communities were maintained until the early modern era (see [section 2.5](#)). Crucially, many of the city-ports mentioned in this book were already established before the arrival of European colonizers, as is shown in the histories of Java (e.g. Banten), Sumatra (Aceh), and the Malay Peninsula (Reid 1993b). This is also confirmed by the trade languages that emerged, in particular the contact Malay varieties discussed in [chapter 3](#).

In these regional centres which developed as a consequence of maritime trade and which must be seen as interethnic and multicultural settlements, we also find social patterns that differ markedly from the inland agricultural socio-economic structures. The former tended to be cosmopolitan and heterogeneous communities with dynamic populations that followed monsoon rhythms, while the latter were typically conservative and more homogeneous, stable societies (Reid 1993a, b; Pearson 2003). Historians tend to describe the city-ports as united by shared cultural beliefs and practices that go beyond ethnic and linguistic ancestral identification. Religious rituals, for example, are typically focussed on celebrating the sea and safe passage, rather than land-based rites typical of the inland. Traditional Muslim and Hindu cultures have in fact been sceptical and timorous of the sea, thus not really encouraging sea trade; however, living in the city-ports meant distancing themselves from orthodox beliefs in order for individuals and communities to engage in such trade. This distinguished land-bound agricultural people from seafaring ones.<sup>3</sup> Marine communities were, and still are, common in the region, as can be seen in the lifestyles of ‘floating villages’ still found – camouflaged as eco-touristic attractions in Southeast Asia – in the swamp-dwellings once common, for

example, in the state of Kerala, and in the economic role played by piracy to this date (see Reid 1988, 1993a; Pearson 2003). The recognition of this sense of unity provided by a shared economic lifestyle, together with the pluralistic ethnic and cultural composition of coastal settlements, is important in order to appreciate the nature of bonding that typifies these networks and offers a fertile ground for the formation of new languages. For one thing, the city-ports were quite removed from the centralized, homogenizing ideology of one nation, one race, one language that Western colonizers would eventually introduce into the region. Furthermore the degree of intercultural exchange across the Indian Ocean region was very high, as is manifest in the nature of lexical borrowing: for example, over five thousand items of Arabic origin can be found in the Malay lexicon (Knappert 1985).

While islands are rare and far apart throughout most of the Indian Ocean, and were usually uninhabited until recently, as is the case for the Cocos (Keeling) Islands (see chapter 3), the Indonesian archipelago stands out as a cluster of smaller and larger islands with a long history and a significant role in the development of the region. Because of its central position within Monsoon Asia, as well as its natural resources, in particular the highly desired spices found especially in the Maluku Islands,<sup>4</sup> the Indonesian archipelago and the Strait of Melaka define much of the history of this region, providing a focal point for Chinese as well as Indian economic and political activity (Reid 1988, 1993b, 1996). Because of this central role, akin to that of the Mediterranean in the west (Reid 1980: 235), the Indonesian archipelago assumes a central position in the historical sketch presented here.

## 2.2 Sino-Japanese trade

Among the dominant powers of precolonial Monsoon Asia, we must certainly consider the Muslim sultanates, India, and China, as well as a number of predominantly Javanese trading empires. In this section, I focus on the Sino-Japanese trade as one of the most influential aspects of the region's history, especially in terms of population movements and interethnic relations. In the historical chronicles of the Ming Dynasty (AD 1368–1644), it reads:

Envoys with proclamations of the Imperial accession and instructions were sent to the countries of Annam, Siam, Java, Ryukyu, Japan, Xi-yang, Samudera and Champa. The Emperor had previously issued instructions to the officials in the Ministry of Rites, saying: “During the reign of the Gao Emperor Tai-zu, all the *fan* countries sent envoys to the Court and they were all treated with sincerity. Those who came with local goods were allowed to trade and had things made easy for them. When there were people who did not know how to control their envy or offended against legal regulations, all were forgiven, as a way of cherishing those from distant lands. Now all in the four seas are one family. It is proper that we widely proclaim that there are no outsiders. All countries which wish to express sincerity by coming to offer tribute are to be allowed to do

so. You should send instructions to these countries so that they are clearly aware of my will.” (*Ming Shi-lu*, vol. 9: 205)

The spice trade, especially the trade in cloves, pepper, nutmeg, camphor, and cinnamon from the Maluku Islands, was by far the single most important factor in the region during precolonial and colonial times. Traders from the Arab Peninsula, India, and China, who could easily sail back and forth seasonally following the monsoon patterns, had been attracted to the spice trade between Maluku, Sumatra, Malaysia, and Sri Lanka from early times. Concrete evidence of the use of cloves can be found in Chinese historiography from as early as the third century BC, and then later in the fourteenth century AD, with specific reference to the Maluku Islands (Reid 1996). An even larger and more active network was to be established by Chinese, Arab, and Javanese traders from the fourteenth century onwards, with Melaka as the principal port and trade centre, as described below. Evidence of the intensity of contacts lies in the existence of the trade language known as Bazaar Malay, characterized as a language of interethnic communication from as early as the seventh century AD, which developed as a consequence of the spice trade and its related activities (Adelaar and Prentice 1996; chapter 3 of this volume).

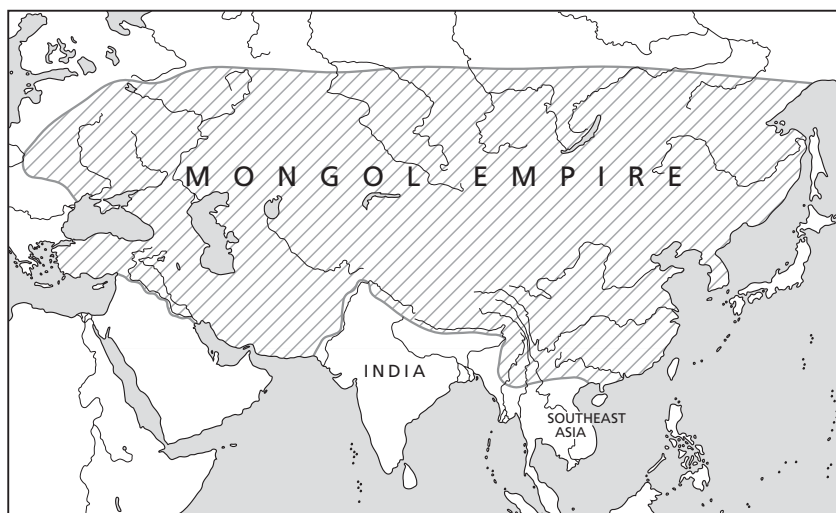
While the rest of this section highlights Southeast Asia’s contacts with China, it must be said that this was not the first contact between a distant empire and Southeast Asia: the Tamil-speaking Chola Dynasty, located on the eastern coast of India, had been trading with Sumatra-based Sri Vijaya between the ninth and thirteenth centuries AD and reached China in the eleventh century (Sastri 1976; Sardesai 1997). Their influence in Southeast Asia is particularly distinctive in cultural aspects, such as the arts, and in literature and writing. Indian influence in general is present in the languages of Indonesia and Malaysia, the clearest being from Sanskrit, which forms the bulk of loanwords in Indonesian and is still a source for many neologisms. There is also phonological influence of Sanskrit on Javanese, Madurese, and Old Malay. The presence of Arabic in the region should also be acknowledged, particularly intensive from the thirteenth century onwards in relation to religion and trade. Arabic lexical influence is consequently heavy in Malay. An example of a mixed language that may be related to Indian influence is Chitty Malay, in which Malay blends with South Indian features (see chapter 3). Likewise, a particular style of literary Malay used for religious texts, known as Kitab Malay, is rich in Arabic syntactic influence (Adelaar 1996a: 695–7).

### 2.2.1 *The Mongol Empire*

In the thirteenth century, the Mongols, who already controlled most of Central Asia, conquered China and founded the Yuan Dynasty. They also occupied and controlled much of Southeast Asia and were responsible for an injection of

Chinese technology into the newly occupied areas, as can be seen, for example, in the hybrid boats called ‘junks’ that are typical of different cultures within the region (Reid 2000). During the Yuan Dynasty (AD 1271–1368), the Mongol Empire, at the peak of its strength, covered an extremely vast area, from northern China to mainland Southeast Asia, and from the South China Sea to the eastern borders of Europe (map 2.2). For as long as it lasted, the Empire provided safe passage on most overland routes between East and West, and promoted trade, not only overland but also by sea; because of this relatively stable situation, the period is referred to as ‘pax mongolica’ or Mongol peace. Not all of the Mongol reign was peaceful, though: in 1293 the Mongols sent a large punitive force of 20,000 soldiers from the province of Fujian, in southeastern China, to Java, in order to submit the Singhasari Kingdom, which was refusing to pay tributes to China (Sardesai 1997). In a bizarre turn of events, the Singhasari prince Vijaya managed to exploit the situation to his advantage and, after expelling the Mongol forces, founded the Majapahit Empire, which was to become the strongest precolonial empire of the Indonesian archipelago. Under this empire, Java rose to become a cosmopolitan centre where merchants from China, Cambodia, Vietnam (then Dai Viet), Thailand (then Siam) and India traded goods – spices, in particular, but also technology and culture (Reid 1988, 1993a).

During the fourteenth century, trade with South China intensified, as attested by Chinese historians, and the relationship between China and Indonesia



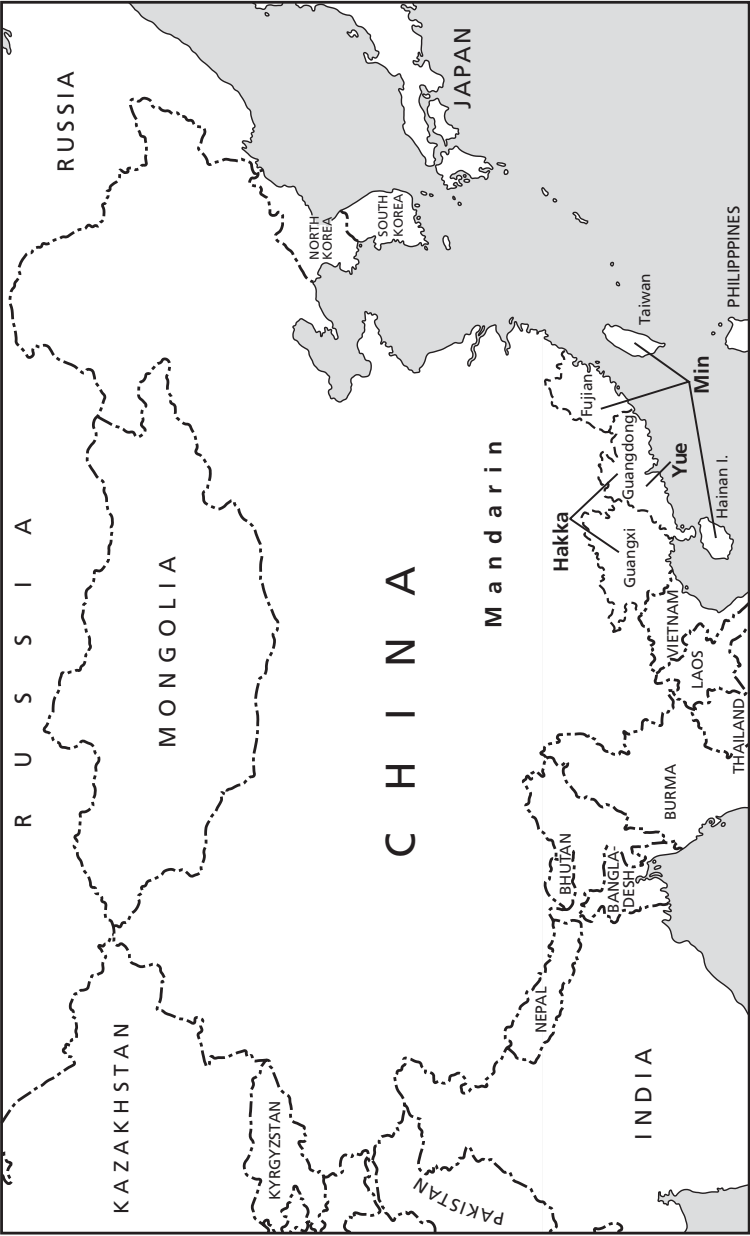
Map 2.2 Mongol Empire *ca.* AD 1300

became particularly tight. There are a number of accounts by travellers and merchants that testify to the involvement of Chinese merchants, predominantly from Guangdong and Fujian, the two most prominent regions in Chinese history when it comes to maritime trade, in the spice trade at this point in time (Galvão 1544). The former, together with the island of Hainan, offers the shortest possible route to mainland Southeast Asia, and has traditionally had strong cultural and historical links with that region (see map 2.3). The latter had two fundamental reasons to concentrate on the sea: (1) it was the launching point for trade with Taiwan and Japan; and (2) for a long time it stood outside the direct dominance of the imperial court, cut off from the hinterland by high mountains. Therefore the Fujian region relied for its economy on its fishing and trading fleets, and its seafarers became the most intrepid travellers of China, as is attested to this day by their large presence in many Southeast Asian countries, especially Indonesia and Malaysia, but also Thailand, Vietnam and the Philippines (Pan 2000: 30).

While Chinese communities had already started to appear in Southeast Asia in the eleventh and twelfth centuries, by the fourteenth century their presence was well established (McPherson 1993: 131). Later in that century, historical sources record a shift from Chinese involvement to Javanese Muslim presence in the trade. Reid (2000) explains that this represents not so much a replacement of ethnic Chinese by ethnic Javanese but rather “an incorporation into a Java-based political project of existing trade networks, in which Chinese and Muslim merchants had played key roles” (Reid 2000: 62). This is an important factor as it attests to the length and intensity of cultural contacts and intermixing of Chinese and Javanese, which led to new cultural identities in particular during the urbanization witnessed in the city-ports (see section 2.3). This shift in identities is supported by two parallel developments: (a) the expansion of the Majapahit Empire based in Java; and (b) the rise of the Ming Dynasty (AD 1368–1644), which replaced the Mongol Yuan Dynasty in China. The latter factor is relevant in this account for two reasons elaborated on in the following sections: (i) the Ming Dynasty changed the nature of commerce from liberal, individual entrepreneurs to large-scale operations under imperial control (before discouraging Southeast Asian commerce altogether); and (ii) the fall of the Mongol Empire brought an end to the safe overland routes and triggered a further expansion of maritime commerce.

### 2.2.2 *Zheng He and the early Chinese Diaspora*

The ban on private commerce introduced by the Ming Dynasty was a significant event in the history of Monsoon Asia. Consequently, Chinese merchants, known as *huasheng*, ‘merchant sojourners’ (Wang 1996), who still wanted to engage in trade were forced to consolidate their bases outside China, leading



Map 2.3 China: coastal regions and languages

to the establishment of more permanent Chinese communities in Southeast Asia. At the same time, the Ming rulers commanded official missions, in the form of the imperial fleet, and sailed to Southeast Asia with economic as well as political mandates. These missions have to be appreciated in an Asian context, as they did not involve small vessels with a few selected ambassadors. As already implied in the numbers of the Mongol punitive force mentioned above (20,000 strong), these were massive affairs with the full backing of imperial structures larger and more advanced than anything the West could even imagine in that period. There were at least seven such missions between 1406 and 1432 under the command of the Chinese Muslim general Zheng He (Reid 1996). Zheng He was a former court eunuch very close to the Emperor Yong Le. Under his supervision, a large fleet was built that, at its height, counted up to three hundred ocean-going vessels capable of carrying as many as 30,000 sailors. The fleet consisted of technologically advanced ships, roughly six times the size of Columbus's caravels, which could transport men, women, soldiers, animals and provisions, as well as gifts for tributary states. Zheng He led his fleet through Southeast Asia to India, Sri Lanka, the Middle East, and as far as Somalia (Mogadishu was reached by the Chinese fleet in 1418). The primary mandate of the fleet was not to pillage and conquer, but to establish new political alliances as well as to research and discover new regions. At the same time, an important part of the expeditions was commerce. On their journeys, each fleet had to station on the northern coast of Java for several months waiting for the easterly monsoon to arrive and carry them towards Melaka. From the chronicles of general Zheng He, we learn of many Javanese ports in which Chinese communities already thrived, such as Palembang, Gresik, and Surabaya.<sup>5</sup> These communities are described as of either Cantonese or Fujianese origins, i.e. from southern and southeastern China (Ma Huan 1433). The arrival of the imperial fleet in the early city-ports of the northern Javanese coast and the Melaka Straits enhanced the financial importance of the existing Chinese communities, as these were the natural point of contact between the fleet and indigenous powers. Moreover, each expedition left behind substantial numbers of people, in the form of men and women traders, who became permanent members of the coastal settlements (*Ming Shi Lu*, vol. 9). The communities' populations were further increased by substantial defections from the soldiers' ranks, sometimes involving thousands of individuals. Such was the intensification of trade that followed these naval expeditions that this period has been called the 'age of commerce' of Southeast Asia (Reid 1993a, 2000), characterized in particular by the development of a number of city-ports.

### 2.2.3 The 'end' of Chinese maritime commerce?

Towards the end of the fifteenth century, a conflation of financial difficulties, possibly caused by the costly maritime expeditions, together with official ideological opposition to political and commercial liberalism, as well as concerns about hostile Mongol tribes in the north, led to a major shift in the history of China. The Emperor turned the interests of the empire away from trade with the Southeast Asian region, ceased the missions between China and Java, and banned maritime commerce altogether. This left the coastal communities of Guangdong and Fujian in an impossible economic situation, used as they were to relying on the sea as their main source of survival and profit. The result was an increased number of merchants from Canton and Fujian fleeing China in order to establish their trade elsewhere (Reid 1993b, 1996, 2000). It is clear that this shift in international politics did not lead to a withdrawal of Chinese traders from the Monsoon region; rather, as argued in Ray (1995), what we see is a shift from state-sponsored trade to private trade.

From the end of the fifteenth century, we see an increase in Chinese people in the coastal settlements of Southeast Asia, as well as a partial isolation of these communities from China. Java-born Chinese communities became progressively indigenized, the result of a gradual mixing of Chinese, Javanese and Muslims that had started at least a century earlier (section 2.2.1 and Reid 1993a, 1996, 2000). As suggested in numerous chronicles and travellers' diaries, the distinction between Chinese and Javanese became blurred in the city-ports, as they often had common origins and common appearances, to the point that even the king of Java, it was rumoured, was a Chinese vassal, not a tributary (Pires 1515; Barros and Couto 1788). The *Ioas* (Javanese) of Melaka described in Portuguese chronicles, and later encountered by the Dutch on the northern coast of Java (e.g. in Banten), were indeed the powerful descendants of the Sino-Javanese communities who inhabited the city-ports and other coastal settlements and who would be known later as *Peranakans* (Pires 1515; Reid 2000: 73).<sup>6</sup> These were Java-born traders who presented themselves as descendants of Chinese and were closer to Chinese people in both appearance and customs. In point of fact, the Malay maritime code of Melaka reflects Southern Min influence (spoken in southern Fujian), e.g. *kheh-ui* 'passenger-space', and Cantonese influence, e.g. *daching* 'weighing scale' (Reid 2000: 75–6). Also from this time comes the word *cengkeh*, meaning 'nail' in Southern Min, borrowed into Malay and Javanese to refer to cloves (Reid 1993b). The intensity of trade between China and the Malay world throughout this period is also attested in a colloquial Chinese–Malay lexicon in which a number of terms for items of trade are registered, dated around the mid-sixteenth century (Collins 1998: 14; chapter 3 of this volume).<sup>7</sup>



#### 2.2.4 *Islam*

As the presence of Muslim traders has been mentioned above, a note on the role of Islam in the region is necessary here. During the period of intensification of trade between China and Java, another development that had been taking place was the spread of Islam. While Arabs had been ubiquitous in the region since the eighth century, it was not until the thirteenth century that Islam spread and became a dominant religion. At this time, Muslim traders stationed in Gujarat (northwest India) moved through India and into the Malay/Indonesian region. A combination of various factors may have led to the success of Islam, in particular the internal fragmentation of Hindu politics, Sufi missionary work, and the reaction against a new aggressive force, Christianity (Sardesai 1997). Between 1402 and 1404, Muslim traders teamed up with Chinese interests and Sumatran factions opposed to the Majapahit Empire, and established the multiethnic port of Melaka, which quickly became the ruling centre of the region (section 2.3). It must be noted that Islam became accepted by different peoples and cultures relatively quickly, partly because it was never imposed, and partly because it presented positive, even encouraging attitudes to travel and commerce, consistent with numerous passages in the Qur'an praising the sea and trade (Parkin 1994; Pearson 2003). However, for the purpose of this chapter – which is to outline those elements that form the ecology in which the contact languages described here evolved – Arab and Muslim history plays a very limited role and will therefore not be discussed in detail. It is clear that in precolonial times, people of Muslim faith were culturally integrated within the commercial culture of the city-ports. 'Polarization' of Muslim and non-Muslim cultures was not part of the early modern history in the region, and it is only in the late sixteenth century, as a consequence of the spread of Christianity and the antagonism between Europeans and Arabs in the spice trade, that the dichotomy between pro- and anti-Islamic peoples arose within Asian communities (Reid 2000).

### 2.3 **The city-ports**

The Sino-Javanese communities introduced in section 2 were not merely settled by two clearly distinct ethnic communities, viz. the Chinese and the Javanese. It has to be noted that, first of all, there were at least two distinct Chinese groups, the Cantonese and the Hokkiens, as the Chinese with origins in southern Fujian are known in the region. Other prominent groups were the Hakka and the Teochew. These groups would have spoken mutually unintelligible languages, with high dialectal diversity within each group (Norman 1988; Ansaldo 1999). Language and dialect diversity applies even more evidently to the peoples from different parts of the Indonesian world. Java, Borneo, Maluku, Sumatra and the Malay Peninsula are inhabited by populations known to be

ethnolinguistically very diverse, and who speak varieties that are not mutually intelligible. In addition, one must factor in the presence of traders of Indian, Arab, and, later, Western provenance. The early city-ports of Java and other trade centres further east, in the Maluku and in mainland Southeast Asia, were therefore characterized by ethnic diversity, multiculturalism, and cosmopolitanism. These ‘urban’ settlements stood in stark contrast to the interior, which was often perceived and described as conservative and even underdeveloped (Ma Huan 1433; Than Tun 1983). Western sources also note how open and welcoming these city-ports were, virtually too diverse in terms of culture and religion to display feelings of superiority towards one another (Marsden 1812). Used as they were to different customs and languages, the Asian rulers expressed no objection to the Westerners’ participation in their trading networks.

It is also important to understand that premodern Southeast Asia, defined as the period from the mid-fifteenth to the mid-nineteenth century AD (Reid 1988), had already started witnessing a real process of urbanization through a rapid growth of commerce and was, by the sixteenth century, one of the most urbanized regions of the world (Reid 1980: 239). The shift from agricultural states to city-based economies was the result of a number of factors. Firstly, it was a continuation of, in particular, three influential trading empires of the first millennium: Funan, Champa, and Sri Vijaya (Reid 1993a: 67). While the first was located on the Mekong region of Mainland Southeast Asia, the other two were based in the Indonesian archipelago. The presence of influential trading empires based in the Indonesian archipelago ensured an early dominance of a Malay-based language as the language of trade. Secondly, the continuing rivalry between Muslim and Hindu powers led the Hindu kingdoms to move eastwards, leaving a number of city-ports of northern Java and the strategic Malay Straits in a relatively independent position. These ports inherited a long tradition of regional, pluralistic meeting points and grew to become large urban centres. By around 1500, Melaka already had, by generous estimates, between 100,000 and 200,000 inhabitants, roughly 20 per cent of the Peninsula’s population (Reid 1993a: 75).<sup>8</sup>

In order to gain a better understanding of the social dynamics of a city-port, below I sketch in some detail the basic structure of Melaka, historically probably the most significant of the city-ports, and a well-documented example of a city-port of early modern Southeast Asia. As such, it offers an invaluable insight into the social and cultural dynamics of a community in which an important contact language flourished, namely, Baba Malay (see chapter 7).

### 2.3.1 *Melaka*

For the study of city-ports in precolonial times, Melaka offers an interesting case on at least two counts: (i) there is good historical documentation, in particular the Malay Annals (*Sejarah Melayu*), which describe 600 years of history of the sultanate, compiled in the seventeenth century, and the account of Tomé Pires (1515), compiled in the early days of Portuguese rule and therefore indicative of life during the Malay sultanate; (ii) Melaka was seen as a model by other city-ports, as indicated by the fact that its legal code was adopted by several other city-ports of the region. Though the historical documents on precolonial life in Melaka are not exhaustive, they offer enough evidence of social life to show the particular propensity to incorporate exogenous patterns typical of city-ports, as well as the cosmopolitanism and multiculturalism that characterized the coastal dwellings of the region.

Melaka (from Arabic *malakat* 'market') is strategically located at the narrowest point of the Strait of Melaka, in an ideal position for guarding and controlling naval passage. It is sheltered by the large mass of Sumatra, which provides protection from the monsoons and makes it a safe harbour for boats to anchor in. Malay and Portuguese sources show that it was the prince of Sri Vijaya (Palembang, Sumatra), Paramesvara, who first established Melaka as a trade centre in 1402, partly convincing and partly coercing boats to call at his port. In its early days, Melaka was inhabited by various people from the Indonesian region as well as by traders from Siam (Thailand). As a testimony to the intimate contacts between China and the Malay world, only a year after the foundation of Melaka, an official Chinese mission arrived from China; a few years later, Melaka's ruler was officially received in China as a vassal. Befriended by China, the city of Melaka gained a reputation for being unassailable, and soon merchants and warehouses from India, China, Java, and Arabia competed with one another in the streets of the city.

Melaka can be seen as the continuation of the economic structure developed by Malay sultanates such as Sri Vijaya, based on trading networks between different entrepôts scattered around insular and mainland Southeast Asia. It was located between two different economic spheres: on the one hand, the pluralistic world of the insular sultanates to the south and east, referred to as 'the lands below the winds', which were rich in spices, and, on the other, the empires of India and China, to its northwest and northeast, which were rich in textiles and other manufactures, and referred to as 'above the winds'. Its trading network extended as far as the Persian Gulf on one side, and Japan and the Philippines on the other. It was a large city, with up to 200,000 inhabitants during the monsoon season (Thomaz 1993).

The city-port can be described as a mercantilist state in which the sultan and the whole ruling class were engaged in trade. Trade was basically the

only reason for Melaka to exist; it depended otherwise almost entirely on its neighbours and vassal states for its day-to-day subsistence. Though basically Muslim, it functioned with two different legal codes, the Malay customary law (*adat*) as a default, and the stricter Qur'anic laws when necessary. Its ruler was a sultan, according to Muslim culture, but the court followed Hindu-Buddhist patterns. For its protection Melaka relied on a mercenary Javanese army; for labour it made use of indentured servants, deployed as sailors, and royal slaves, who managed the trade. It is likely that there were in Melaka around 3,000 public slaves and 3,500 private slaves. The relationship between the state and the citizens was one of mutual collaboration: the state protected its community of traders with its fleet and its political alliances, and the traders in turn administered the city's commerce and provided slaves in times of need, such as war (Manguin 1983; Sardesai 1997).

There were four principal communities in Melaka: Chinese, Javanese, Tamil and Gujarati. Next to these, Armenians, Jews and Arabs also lived in the city. Pires counted at least 6,000 traders and reports that "in the port of Malacca very often eighty-four languages have been found spoken, every one distinct, as the inhabitants of Malacca affirm" (1515: 268). It is reported that, in 1501, five hundred different moneychangers could be found in just one street. All kinds of goods were available in the city: from the most precious spices to sandalwood, pearls, silk, and the feathers of the bird of paradise from Papua New Guinea. Trade was administered through four official harbourmasters, one for each of the four dominant communities. These harbourmasters may also have had a military role in times of war, in particular the duty to gather servants and organize them into troops for the city's army. Ships from 'above the winds' calling at Melaka had to pay 6 per cent import duty to the state, while ships from 'below the winds' paid no duty but had to sell 20 per cent of their cargo to the state at 20 per cent below market price (Thomaz 1993). There is no evidence of discriminating taxation practices based on religion or ethnicity since all merchants were encouraged to reside in the city by the sultan. In fact anyone could become a merchant by free choice, and there were no ethnic groups not welcome by the state. Traders who did not own boats would invest their money with parts of ocean-going cargoes; if the trip was successful, they would receive part of the profit after having paid a commission to the ship-owner. As residents of Melaka, they had the benefit of reduced import duties amounting to half of the standard 6 per cent. Chinese ships did not pay duty according to Pires's account (1515: 274), but presented the court with generous gifts – a sign of the prestige accorded to China and the Chinese.

The revenue from import duty constituted the bulk of the state's economy. Melaka's commercial laws were heavily influenced by Islamic law, which addresses issues of debt and bankruptcy, and explicitly forbids usury. Instead, it

recommends to parties in commerce to agree beforehand on sharing the profit, and determines that, if the capital is lost, no compensation shall be required, unless negligence is involved. This commercial philosophy, we will see, stood in stark contrast to the aggressive exploitation introduced to the region later by the Portuguese and the Dutch.

The extensive degree of urbanization seen in Melaka, as well as other city-ports of the region, in this historical phase, can be accounted for in environmental terms (Reid 1993a: 77). Besides the confluence of historical factors presented above, we must also note that a rice-based economy requires less effort to produce substantial surplus, when compared to barley and wheat. Based on seventeenth-century estimates (in Reid 1993a: 77), we know that, before the Portuguese occupation, between 175 and 195 rice cargoes served Melaka each year from Burma, Siam, Java and Coromandel. Calculating an average of at least 50 tons per vessel, this would have been enough rice to feed 50,000, enough for the more conservative estimates for Melaka's population of the time. Further, as the city-ports were strategically positioned on waterways, they benefited from low-cost transportation. Finally, the city-ports were not parasitic economies but participated actively in the region's economy, attracting trade and paying for their subsistence on the open markets (Reid 1993a).

### 2.3.2 *Merchants*

The most dynamic and influential class of people in city-ports such as Melaka was undoubtedly the merchants, referred to as *orang kaya* 'rich men' in Malay, a term that is also an aristocratic title (Reid 1993a: 115). This account follows a detailed study in Reid (1993a, b), according to which it was possible to distinguish three different groups of merchants: (i) foreign merchants, (ii) semi-foreign or assimilated mediators between the court and the traders, and (iii) local aristocrats engaged in trade.

Foreign merchants were in close relation with royal families and aristocrats, with whom they engaged in commerce. It is in fact not always easy to distinguish between the group of merchants and the group of aristocrats. Although status was an important element in the city-ports, and display of wealth and manpower was also common, the boundaries between the two groups were very fuzzy, and there was a lot of mobility across status lines (Reid 1993a: 122–4). On the one hand, many aristocrats had started off as merchants; on the other hand, many royal families moved to the city-ports to become traders. Also, it was possible for foreign merchants to join the local aristocracy through alliances, marriage, conversion, and so on within a generation. When commerce started to lose momentum, wealthy merchants turned to investing in

land and people. Reid (1993a: 123–4) also notes that in times of rapid commercial expansion, foreign merchants were often chosen as rulers. Being foreign could be an advantage as it meant the availability of international networks as well as a certain degree of impartiality to local affairs.<sup>9</sup> This made the foreign *orangkaya* an influential class, regarded as an asset, not a threat, though its relation with the ruling class was always in a tension between collaboration and competition.<sup>10</sup>

A particular type of merchant entirely associated with overseas trade was the travelling merchant, known as the *huasheng* in Chinese. These were permanent foreigners of very high status, as can be seen in the prominent role they had in the drafting of parts of the Malay Code of Maritime Law (*undang-undang laut*). Five commanders (*nakhoda*) of maritime traders were involved in writing the code, in which *nakhoda* are given power over life and death onboard the ship and the right to sell their merchandise before any other merchant onboard upon arrival in a port. Colonial chronicles give much attention to this group: they note that travelling merchants were particularly welcomed by aristocrats and royal families. It also appears that Chinese and Malay groups were the ones most closely associated with or identified as travelling merchants, a fact confirmed by Chinese historians (Wang 1996), and, along with other travelling merchants, spoke Malay as a first or second language (Reid 1993a: 126). When Melaka was at its highest point of expansion, the elite of native merchants were either Javanese or Sino-Javanese and spoke Malay. After the arrival of the Portuguese, many travelling merchants migrated, leading to the emergence of diasporic societies throughout the region. At the same time, Sino-Javanese presence survived Western occupation: Portuguese sources document the existence of a *kampung Cina* in Melaka, that is, a Chinese quarter; and Dutch sources report the presence of Batavia (today Jakarta, Indonesia) Chinese in the city.

### 2.3.3 Summary

As has been made clear in the preceding account, Melaka was a pluralistic state in which cultural systems of different origins mixed in a pragmatic way: Islam was pragmatically adopted at the official level, and the law was adapted so as not to inhibit trade. For the same reasons, taxes were low, and the state did not control but rather participated in the economy of the sultanate. It was a multicultural state where different groups were well represented and shared power in a quasi-egalitarian way; mobility between social classes was high. In precolonial times, slavery accounted for only a small part of the economy (see section 2.4). Geopolitically, Melaka was in relative harmony with its neighbours, offering a share of its wealth in exchange for provisions and manpower, though of course conflicts occurred. It is interesting to note that Melaka, like most other city-ports of the region, was not fortified; it was not common for

Asian city-ports to erect city walls, a practice which only developed in the sixteenth and seventeenth centuries after the arrival of the Portuguese and the Dutch (Reid 1993a: 88).

The brief sketch provided above serves to highlight the fact that, just like Melaka, city-ports throughout the region were indeed the product of a combination of different political and cultural systems in precolonial times. In this sense, though subject to centralized rule, they were places in which everything from legal systems to religious beliefs was negotiable and negotiated, until the recombination of aspects of originally different systems yielded the optimal system to suit the new, emerging social structures. Similarly, these city-ports were linguistically extremely diverse, considering the number of different languages found in such environments. As is typical of linguistically diverse ecologies, individual and societal multilingualism was frequent, and languages of interethnic communication, in the form of pidgins or any other *lingua francas*, abounded (see [chapter 3](#)). In some cases, such as Bazaar Malay, they have survived to this day (Khin Khin Aye 2005). Before we move on to the linguistic significance of this history, however, one more salient feature of the region needs to be discussed: the nature of human resources in precolonial and colonial Southeast Asia.

## 2.4 Manpower in early Southeast Asia

Marriage and slavery are discussed in some detail in what follows because patterns of exogamy–endogamy and dynamics of manpower are interesting for contact language formation, in the sense that they can be important factors in the formation of new communities. Inter marriages between different linguistic groups, for example, may result in the emergence of mixed languages (Matras and Bakker 2003). Slavery, which often implies substantial population movements, can result in the establishment of diasporic societies, again fertile ground for the evolution of new languages, as is obvious from genetic creolistics. Moreover, investigating these domains reveals important aspects of social relations and allows us to better reconstruct the likely social environments in which new communities and languages developed in this region. Within these two domains, some general aspects that struck Western travelers and historians as markedly different in the Southeast Asian region were (a) the relative power of women and (b) the independence of ‘slaves’ and the relatively mild treatment they received by their owners. Let us examine these aspects in turn.



### 2.4.1 *The role of women*

The social and economic power held by women is often remarked on in accounts by early Western travellers (see e.g. Reid 1996). In Siam, for example, most trade passed through the hands of women (Rabibhadana 1969). In the Malay/Indonesian region, local women clearly enjoyed a more egalitarian treatment than their European counterparts. For example, they participated in everyday commercial activities, and could be found travelling alongside men on Malay and Javanese boats (Reid 1983, 1996; Sardesai 1997). In precolonial Southeast Asia, women could be found in royal courts as attendants as well as armed guards; the female members of royal families engaged in trade just as their male counterparts and the sultan. Considering the high numbers of travelling merchants, in particular the Chinese *huasheng*, temporary marriage for pragmatic reasons was a relatively common practice in the city-ports. Travelling merchants were looked upon as prestigious elements of society, and it was common for them to be married to local women, who were looked upon as good business partners, capable of handling small trade and money changing and therefore a valuable resource. This made women financially independent and allowed them to be in control of their networks of relations. Besides offering local knowledge, women stayed behind to man the business while the men were away on trade during the monsoon season (see Reid 1996, 2000). Short-term marriage was particularly frequent among Chinese traders, but so too was divorce, which was facilitated by law. The practice extended to early Western arrivals too (and perhaps continued even later). After the arrival of the Portuguese in the region, some of their commanders were offered local wives on Java and in the Maluku (Pigafetta 1524; Albuquerque 1557; Barros and Couto 1788; Bausani 1972).

Another widespread practice in the region was concubinage, especially common among wealthy men and members of royal families. Having numerous female partners (just like having numerous servants, see below) was regarded as an indication of status. Moreover, it provided upward mobility to women from poor social backgrounds who could become prosperous and respected (Warren 1981; Reid 2000). The practice of males acquiring bonded women as housekeepers, assistants, and lovers was particularly pronounced in the Chinese communities, and Chinese traders were known to have many wives with whom they had many children (Scott 1606 in Foster 1944; Dampier 1699). These women could be sold if the males returned to their home country but, as the males often became rooted in the city-ports, the women would become their wives. There is therefore a continuum between the practice of concubinage and interethnic marriage patterns in the region. It is important to note that female servants were in general not regarded as prostitutes in local cultures. And while servant–master relations were common, this only applied



to unmarried women, as it was either forbidden or very dangerous to have relations with married servants (Rabibhadana 1969). Western observers also appreciated the role of local women and commented on the higher numbers of women in respect to men found in the community, as the former were destined for the ‘marriage market’ (Reid 2000). Indeed, by the end of the eighteenth century, a large population of *mestiços*, people of mixed blood with usually a Portuguese (or other European) father, could be found in many ports of the region (Pearson 2003).

In point of fact, Skinner (1996) identifies the patterns of marriage described above as the main factor for the development of ‘creolized’ societies in Southeast Asia. Given the fact that Chinese immigrants into the region were predominantly males and frequently took local wives, this is unsurprising. But, crucially, Skinner points out that the main reason why the new communities emerged was the fact that the offspring of such unions were usually not incorporated in indigenous societies. This is partly because Chinese males often married women of mixed ethnic heritage and their children tended to stay and identify themselves within this intermediate social sphere rather than with the more indigenous one. As we will see in the case of the Baba and the Sri Lanka Malays, in these social spheres, endogamy became the more common practice (chapters 6 and 7). Moreover, these communities continued to develop during and after the establishment of European governance. As we will see in section 2.5, the European impact not only created new opportunities for traders and triggered a substantial increase in immigration of Chinese, Malay, and Indonesian people; it also brought with it an active policy of interethnic division and racial classification that contributed to the isolation, and therefore the formation, of new communities in the colonized territories. As we will clearly see in the formation of Baba Malay and Makista (chapter 7), intermarriage alone is not always a sufficient condition for the evolution of new languages. It is only when exogamy is replaced by endogamy that real focussing happens, and it is only then that stabilization of the new variety occurs.

#### 2.4.2 *Bondage*

Whether slavery should be regarded as a universal, cross-cultural concept or whether different cultures in their time produce different interpretations of slavery is an important question that can be answered by examining slavery practices in different cultural and historical contexts. The Southeast Asian region has often been taken to provide us with examples of relatively mild slavery compared to the inhuman conditions the Atlantic slave trade is mostly associated with (Dampier 1699; Montesquieu 1949).<sup>11</sup> This can be explained by the fact that, as we will see below, different forms of ‘slavery’ existed in the region before the arrival of the Europeans, and ‘slaves’ had different degrees

of freedom, as well as rights, depending on the category they belonged to. Reid (2000) offers a particularly enlightened synthesis of patterns of slavery to be found in the Southeast Asian region before the arrival of Western powers: what follows relies on that research unless otherwise indicated. It is particularly fruitful to look at the transition from regional, rural practices to those found in the city-ports before and after the arrival of Western powers. This gives us a good indication of the social dynamics of population structure and population movements in the city-ports.

Definitions of 'slave' usually include (a) the state of being under the control of another person and (b) work done under harsh conditions for little or no pay.<sup>12</sup> Niebor's (1919) comparative study of systems of slavery across cultures identifies three common aspects of the definition of 'slave':

1. A human being seen as property of another human;
2. A person who is on a lower level than the mass;
3. A person who performs compulsory work.

These aspects are broad enough to allow different interpretations in different cultures; it is in the observation of how each of these aspects is realized that we can see clear differences between slavery in the Asian region discussed here and modern Western notions of slavery. The key to understanding slavery in this part of the world lies in recognizing three aspects of Asian society at the time: (i) the general hierarchical nature of Asian societies, (ii) the appreciation of the role of 'manpower' in assessing status, and (iii) the dynamics of debt in the social economy of the time.

#### 2.4.2.1 *Hierarchy, debt and capital*

In terms of social organization, Asian societies have been characterized as complex, sophisticated hierarchical systems. For instance, as has been noted in linguistic and anthropological studies, Asian languages typically display elaborate honorific systems in their grammatical and lexical systems (e.g. pronouns in Malay, Thai, Vietnamese) and/or differential grammatical means that identify the status relation between the participants in a conversation (e.g. Javanese, Korean, Japanese). This means that hierarchy is clearly displayed in forms of address, and power relations between participants must be linguistically coded through correct choice of the appropriate grammatical pattern (Goddard 2005). Vertical bonding is a widely accepted aspect of society and the family, where filial respect is expected on the part of the elders and granted on the part of the junior members in most Asian societies to this day (notwithstanding the 'negative' influence of Western cultural values). This vertical bonding extends beyond the family; this is particularly true of premodern times where extended family networks virtually resembled trading companies or 'smaller states' (Dumont 1970). In the indigenous systems of the region, there was thus a sort of magical bond between master and servant that was regarded as a

fundamental aspect of everyday life and was celebrated in literature, the arts, and folklore (Reid 2000: 187). The relationship between master and servant in Asian contexts of the time has been thus interpreted as an extension of the inherent hierarchical nature of these societies.

Another aspect that is relevant for our understanding of the role of servants in Southeast Asian societies is the nature of capital in premodern Southeast Asia. As Reid (2000) reminds us, the notion of land was not particularly relevant; land was abundant in the region and was not really an issue in the city-ports, whose focus was on trade. The population of Southeast Asia around the sixteenth century was very small compared to the wealth of the region, comprising approximately 15 million, compared to the 100 million in Europe (Reid 1983: 157). Rather than large political configurations, the region was characterized by political diversity and pluralistic states. One way in which power and wealth could be measured was therefore in terms of control of men, through credit, interest, or other obligations (Reid 1983, 1993a: 92). As we learn from early Western descriptions, rich men and royals would surround themselves with large numbers of male and female servants, for protection, labour, entertainment, and simple display (see Reid 1983: 157–8). These descriptions and the sketches that accompany them in fact often point out how difficult it is to tell the slaves from the masters, since they could be seen walking next to one another, wearing similar types of attire (typically a *sarong* and head gear) and were otherwise bare chested. We may also be surprised to learn that ‘slaves’ were allowed to keep their own servants, for example, Dampier (1699: 98) describes people of Aceh as all “slaves to one another”. This is an attestation to the widespread pattern of bondage, which was a result of a person indebting their own life in exchange for money. We see, therefore, that it was “fundamental to Southeast Asian social structure that debt implied obligation, and in particular the obligation to labor . . . The object of creditors within this system was not so much to increase their capital by earning interests as to increase their status and power by acquiring dependants” (Reid 1993a: 108).

From the discussion above we can derive that the term ‘bonded servant’ may be more appropriate than ‘slave’ to describe the prevalent pattern of master–servant relation in precolonial Southeast Asia. The notion of ‘slave’ in this region seems to apply only to a restricted set of people, namely war captives and criminals, whose treatment was very harsh (Reid 2000, see also Hoadley’s class of slaves below).<sup>13</sup> Bonded servants on the other hand were found in abundance. Such a class of bonded servants, whose duty it was to work in order to pay a debt, is one of the three broad classes of service, forming the largest source of labour in typical indigenous systems of manpower of the region (Hoadley 1983).<sup>14</sup> The relationship between bondage and status in this part of Asia can be seen in the notion of ‘debt’ as a viable means of being provided for and for availing oneself of upward mobility. Reid (1993a, 2000) places the origin of this system of labour in the role of debt and obligation found in the local

cultures of the region pre-dating the city-port phase. Diverse cultural practices, such as marriage, funerals, commercial ventures and gambling, provided opportunities for people of low social status to incur debt. In order to afford the costs of such practices, men would offer themselves as slaves, i.e. work not for a wage but rather for food (and pocket money). Bondage could be advantageous for an individual since being bonded meant receiving financial as well as social support; in other words, it was a way upwards in society. Besides money, an indebted servant expected to receive a certain number of benefits with increasing seniority. In particular, the master was required to marry his servants at the right time, and give them their own residence. In addition, because members of this class, though never expected to fully repay their master, could nevertheless be expected to pay back a part of their debt, it was in the master's interest to allow them some financial independence such as engaging in their own trade on the side (Scott 1606 in Foster 1944; Sutherland 1983).

Hoadley also distinguishes another class of service comprising a group of slaves, numerically very small, consisting of criminals and prisoners of war, who could redeem themselves and move to the class of (permanently) indebted servants. Marriage between slaves and non-slaves was possible, though the children of the married servants usually belonged to the master. At the death of the master, slaves could even be freed.

Early European descriptions confirm the more benign conditions applying to what they describe as 'slaves' of this region in contrast with those of the Atlantic. They note how relatively well off servants were in local cultures, making them often comparable to, if not better off than, servants in Europe (e.g. Pigafetta 1524; Raffles 1817). This is not meant to suggest a positive image of slavery in the region, as servants were nonetheless bought and forced to work, with the most unpleasant duties and probably the harshest treatment typically assigned to junior members. Nonetheless, three aspects of bonded servants set them aside from the prototypical Western idea of slave: (i) relative upward mobility, (ii) a certain degree of independence, and (iii) lack of clear segregation between master and servant. These clearly had implications for the evolution of the city-ports.

#### 2.4.2.2 *Manpower in the city-ports*

The city-ports developed over time, from small settlements to large, cosmopolitan commercial centres relying heavily on the trade of bonded people. As outlined in section 2.4.2.1, bonded servants were not only a status symbol of the master but also the support structure for the master's extended family and the commercial activities it engaged in. In these urban centres, bonded servants were trusted to sail boats, engage in trade, serve as messengers, run the household, and provide musical and theatrical entertainment. Upward mobility could be seen in a family from the second generation onwards, where the

children of servants enjoyed a level of intimacy and an access to power that blurred the lines between the bonded and the free (Reid 2000). Moreover, mobility was further enhanced by the fact that the dense and pluralistic social setting of the city-port allowed slaves the possibility of leaving (fleeing) a family in case of dissatisfaction and easily finding employment and protection in another household without leaving the city. This meant that a number of laws had to be put in place in order to protect as well as control the movements of servants (Sutherland 1983). It has been noted how employment of private slaves tends to peak at times of maximum development of urban centres engaged in international trade (Reid 1980). Thus, when Western powers entered the trade in the region, they also influenced the slave trade. City-ports such as Melaka, Banten and Makassar could sustain a considerable traffic of people between different provinces, and in time manpower became one of the goods to be traded in the city-ports (section 2.5). But crucial to this account is the realization that until that point, servants and slaves had been integrated in the social scene. Even during Western colonial rule, servants retained a level of intimacy and were extended a degree of respect that makes the comparison with slavery in the Atlantic world very difficult. In this sense, once the diverse populations of servants started being moved around within Southeast Asia and beyond, we should expect aspects of the cultures, in terms of social and linguistic practices, to filter through to the new environments in which they were to participate. In other words, when servants and labourers are mentioned in this history they are considered to have been integrated within the family structure of their employers, and therefore to have played a part in the newly emerging ecologies.

### 2.4.3 *Final remarks*

From the above we can see how bonded servants, slaves, and concubines constituted a fundamental part of the economy and the social life of the region, in particular in the city-ports. In these environments, these people not only supported the local economies, but also fostered a dynamic, pluralistic environment where cross-cultural values and new systems of beliefs were negotiated and promoted. The areas of interethnic contact (and interethnic communication) were many, far exceeding the simple picture of master–servant communication. For a socio-historically realistic scenario, we must conceive of many sub-ecologies such as the market, the docks, illegal shops, brothels, alcohol dens, the impromptu nurseries, and many other licit and illicit environments that constituted the ‘co-habitation societies’ (Faraclas, Corum *et al.* 2007) in which important linguistic and cultural transmission would have taken place. This high degree of integration reminds us at a general level of Chaudenson’s ‘homestead’ phase (Chaudenson 2001), i.e. the early stages of formation of

French Creoles, where Whites and Blacks were in intimate contact. The parallel can be extended to Chaudenson's observations regarding the widespread practice of interethnic unions and marriages found, for example, on Martinique and Bourbon (now called Réunion). Compared to the Indian Ocean settings described in Chaudenson (2001: 99), however, we can note two differences. Firstly, in Southeast Asia there seems to have been a lower degree of integration, or cultural accommodation, of females towards their masters, as women brought with them significant elements of their culture that would influence their offspring (Taylor 1983; Reid 1988: 146; [chapter 6](#) this volume). Secondly, the demographics of Asian environments are different, in the sense that servants often outnumbered their masters in all environments, as was the case in the Caribbean. Therefore, whether or not we subscribe to Chaudenson's claim of the role of Whites in the 'creolization' process of French Creoles in the Indian Ocean, in Southeast Asian ecologies in particular it is obvious that women and servants played a very influential cultural role from the very beginning. The linguistic histories to follow in the next chapters will add some perspective to the role of these people in the evolution of new cultural forms – languages.

## 2.5 The Western impact

In this section, we briefly look at the immediate consequences of the arrival of Western powers in the region. The focus here remains Southeast Asia, as the cradle for many of the social dynamics that typify this part of the world, and the arrival of, in particular, the Portuguese and the Dutch. More specific historical details pertaining to South Asia, in particular Sri Lanka, are provided in [chapter 6](#), while British colonial history in Southeast Asia and the South China Sea is discussed in [chapter 8](#).

We can distinguish three phases of Western influence in the colonial history of Southeast Asia. (i) The Portuguese phase, spanning from the early sixteenth century to the seventeenth century. In this phase, the numerical presence of Westerners in Asia is relatively low. Slaves of African origin are introduced in the region and intermarriages between Western males and Asian women are numerous. (ii) The Dutch phase, starting in the seventeenth century, during which a more organized Western presence enters the South and Southeast Asian region, resulting in the Dutch monopoly of the East–West trade. The Dutch often use the Portuguese as a source of knowledge on local trade and culture and even avail themselves of Portuguese-based trade languages. Slavery and deportation increase, in particular in and from Java. (iii) The British phase from the early nineteenth century to the recent past. In this phase, the East India Company asserts itself as the unchallenged trading power of many parts of the Asian region, often managing other Western nations' interests, as in the case of Southern China (see [chapter 8](#)).<sup>15</sup>

### 2.5.1 *The Portuguese in Asia*

From the perspective of Asian cultures, sixteenth-century Portugal can be seen as a relatively poor and hungry nation fomented by religious fanaticism. Its rulers had two major goals: trade, and conversion of infidels. In 1511, Alfonso de Albuquerque led a fleet against Melaka as part of an attempt to expand its control from India (*Estado de India*) to Southeast Asia and beyond. This would secure a hold on the spice trade and allow Portugal to break the Muslim networks of the region. After failing to capture the city twice, the Portuguese turned to burning Arab and Indian Muslim warehouses, sparing the Chinese and non-Muslim ones. This created a division within the city's communities that eventually led to the departure of the sultan and his mercenary troops. Consequently, the Portuguese were able to enter the city-port, and proceeded to slaughter all Muslim men, women, and children. In a speech recorded after the fall of Melaka, Albuquerque praised the destruction of Muslim trade as valuable to the Lord and beneficial to Portugal's economy (Panikkar 1953).

### 2.5.2 *The Portuguese in Melaka*

The Portuguese established themselves in Melaka for a short period and used this as a base for further expansion east, in particular to Macau (see [chapter 6](#)). They did not however leave a marked influence in the region for a number of reasons. Firstly, their main interest was the spice trade, involving the control of nutmeg, clove, and pepper. Therefore, with the exception of a number of forts and ports on the route to Maluku, the Portuguese never established a territorial presence, confining themselves to a number of coastal settlements (Sardesai 1997). The agricultural hinterland remained fully under indigenous control, and pre-existing powers were not affected. Secondly, as noted above, the Portuguese linked religion with trade and politics, engaging in what were perceived as acts of piracy against Arab and Indian Muslim ships and cargoes. This was in stark contrast to the tolerant religious culture of the region, especially the city-ports, and it alienated the Portuguese powers from the local ones. Indeed, their violent attempt to convert local groups to Christianity has been identified as one of the main reasons for the spread of Islam during the sixteenth century. Thirdly, the Portuguese developed a nepotistic network of port captains, resulting in arbitrary and exaggerated toll charging which undermined trade, further alienating the Portuguese from the local powers (Sardesai 1997).

In spite of the above, the type of colonization practised by the Portuguese resulted in a Portuguese–Asian Diaspora still present in the region (see [chapter 3](#)). This is because Portuguese settlements in Asia were characterized by a crucial difference from other Western enclaves: in their Indian and Southeast



Asian territories, there was a higher degree of interethnic admixture thanks to an open policy of intermarriage and conversion to Christianity encouraged by the Portuguese Crown. This is clear in particular in India, evidenced by the Portuguese enclaves still alive today, such as Diu, Daman and Goa, in contrast to British India where mixed families never became ‘communities’. It has been suggested that this may be explained by the fact that the British always had the manpower and the technology to extend their political control in India without relying on a stratum of mixed-race people (Bethencourt 2005). The Portuguese, on the other hand, developed their empire in spite of the crucial absence of human capital, and depended from the very beginning on local manpower. They were completely reliant on trade, as the option of establishing themselves as landowners was unavailable to them. Portuguese traders, moreover, were predominantly engaged in inter-regional trade and regional markets, as revealed by the activities of the Portuguese merchant fleet in Macau (Souza 2005: 358). They acted as suppliers and intermediaries of Asian goods, while the Dutch and the British were mostly focussed on the global, long-distance trade between Asia and Europe. Through these activities, the Portuguese left behind a higher degree of cultural legacy than the Dutch, as can be seen in the numerous contact varieties in which Portuguese features are salient (e.g. Papia Kristang, Makista) vs. the lack of linguistic influence from Dutch. This can be explained by the fact that the Portuguese were the first Western traders to come into contact with Asian powers in Sri Lanka (see chapter 5), Indonesia (see chapter 3) and Southern China (see chapter 8), a fact that put them in a position of mediators between later Western arrivals and local authorities, as seen in the history of Macau (chapter 8). This meant that varieties of Portuguese were often used by Western traders in their early contacts with Asian merchants (Ostler 2005).

### 2.5.3 *The Dutch in Melaka and Indonesia*

The Dutch entered the region as a far more organized and numerous force under the banner of the Dutch East India Company (*Vereenigde Oostindische Compagnie*, VOC). In 1641, they took Melaka from the Portuguese and gradually extended their control over the trade with Java and Maluku. The Dutch displayed what has been identified as an obsessive concern with profit and an indifference to cultural and political aspects of the region.<sup>16</sup> They engaged predominantly in regulated taxation of goods and, in some cases, controlled cultivation of crops (Pearson 2003). Because they left local political and administrative structures in place to work for them, they had a minimal overall cultural influence in the region. They did however make their influence felt in heavily taxing the Asian–European trade; by the mid-seventeenth century, through a combination of well-organized networks and substantial military enforcement, they had also gained control of the regional networks (Pearson 2003: 150). This was to



affect the local economy, particularly the Indonesian one, leading to the closure of a number of local routes and forcing numerous coastal networks to relocate to the economic systems of the inland. Again, the tax-based system of the Dutch was in stark contrast to regional trading patterns heavily influenced by Muslim culture: Islamic law does not support the idea of ‘interest’, preferring instead a profit-sharing between creditor and debtor (Reid 2000 and [section 2.3.1](#) above).

Another practice introduced during Dutch occupation was a harsher treatment of servants that negatively affected the slave culture of the region. Historians are not unanimous as to which people were more violent, the Portuguese or the Dutch (see Pearson 2003: 122–32). If the former seemed to be more prone to random or religiously motivated violence, the latter were, even in this respect, more systematic. Besides resorting to violence in cases of local resistance towards the tax system or in cases of local alliances with other European powers (Sardesai 1997), the Dutch practised systematic deportation of people from the Indonesian archipelago to other colonies such as Sri Lanka and the Cape (Taylor 1983; [chapter 6](#) this volume). These deportees could be slaves, criminals, or political exiles, usually involving members of royal courts opposing the Dutch, as well as warriors, who would become mercenary troops to fight foreign kingdoms on the side of the Dutch. But deportation was also used to acquire valuable land: between 1619 and 1623, for example, the Dutch deported most of the population of the Banda islands, resettling it with Dutch people and slaves from faraway colonies (Pearson 2003). This trend, later continued – in fact intensified – by the British, led to the creation of ethnically Malay/Indonesian Diasporas in places as remote as Sri Lanka, the Cocos (Keeling) Islands, as well as the Cape in South Africa. By the mid-eighteenth century, the Dutch had virtually gained control of key areas in the spice trade. Moreover, the VOC was the only Western power that Japan would trade with at this point in time, which meant a dominant position in the Asia–Europe trade as well as in parts of the inter-Asia trade. However, the strict organization of the VOC, the vast human resources required, and the costs of supervising and policing the monopoly it had established eventually led the company to bankruptcy (Pearson 2003).

#### 2.5.4 *Western treatment of slaves*

As already suggested in the previous section, after the arrival of the Portuguese and the Dutch, the patterns of manpower population structure described in [section 2.4](#) started to change: in this phase we see the shift from a predominantly bondage-based system to a harsher system akin to slavery as witnessed in the Atlantic trade. In particular, Western treatments of slaves differed from local custom in at least three respects: (i) there was more segregation between Westerners and local populations, (ii) the slaves received comparatively harsher treatment

and frequent punishment, and (iii) in the beginning Western masters tried to hire slaves from faraway regions, rather than locally (e.g. India, Madagascar, New Guinea). Reid (2000) corroborates the fact that Asian and Western treatments of servants and slaves were markedly different when he suggests that perhaps the presence of Indonesian slaves in the Dutch colony of the Cape might have been a reason for the relatively mild conditions that seem to have been obtained there, compared to the New World colonies.<sup>17</sup> In Batavia and the Maluku, the Dutch reportedly executed and tortured slaves rather frequently, and Dutch wives were known to have reserved particularly cruel forms of torture for attractive female slaves who were in a position to cause jealousy (Stavorinus 1798). Deportation and selling to third parties, quite anomalous practices until then, became the norm. The majority of slaves were women, usually taken from small, poor, and fragmented states in eastern Indonesia (e.g. Bali, Borneo, Makasar, Papua) and moved to the west, including the Malaysian Peninsula. While the main centre of 'slavery' in the region during the fifteenth and sixteenth centuries was Melaka (see section 2.3), trading with Java as well as Sulawesi and the Maluku (Pires 1515), by the end of the eighteenth century Batavia stood at the centre of the slave trade that reached as far as the Philippines (Warren 1981).

In 1807 the British made slavery officially illegal, with the Dutch following in 1860 (see Reid 2000). However, it was not until the beginning of the twentieth century that both countries effectively stopped slavery in the region.<sup>18</sup> In addition to slavery, we have to consider the role of (a) indentured labour in the seventeenth and eighteenth centuries, and (b) contract labour in the nineteenth century. Both systems involve the movement of people for the purpose of labour under restrictive 'contracts' of employment for a fixed period of time, followed by repatriation (see the case of the Cocos Islands, chapters 3 and 7). The transition from a small-scale, geographically confined and culturally accepted human trade to a large operation involving substantial displacement of people would be felt in the region in particular in terms of increased ethnic and social segregation and the emergence of numerous diasporas (see chapters 6 to 8).

### 2.5.5 *Rethinking Western influence in Asia*

In commenting on Western trading companies in Asia, Lombard (1981: 187) says:

When seen from Europe, they doubtless appear to be autonomous institutions of wonderful efficacy, heralding the colonial empires of the nineteenth century. When seen from Asia, they seem first and foremost to be uncertain attempts on the part of newcomers to find their way as best as they could into a system which had been in existence for centuries.

Historians of the Southeast Asian region have identified a “disproportionate emphasis on the role of the miniscule body of Europeans” in premodern Southeast Asia (Sardesai 1997: 63), and have called for more attention to the development of regional powers and their networks, even if historical documentation may be limited (see van Leur 1955). Note that the Indian Ocean had seen ‘foreign’ presence since the Romans and Greeks, and Chinese and Arab fleets had been present in the region for many centuries (see Hall 1985: 26–38). Compared to the fleets of Zheng He described above, or to the empire of Suleyman, who from 1520 onwards controlled the coast between Crimea and Yemen, Portuguese and Dutch fleets certainly appeared small and unimpressive to the Muslim sultanates of Southeast Asia (Pearson 2003).

The Portuguese and the Dutch had a limited overall influence on the development of the early modern history of the region. They were small in number, confined to coastal or insular settlements typically isolated from the hinterland, unlike the coastal settlements under Asian rulers which maintained open relations with inland powers (Sardesai 1997; Pearson 2003). One reason for their confinement was that, in terms of government, the Portuguese and the Dutch conflated politics with commerce in a way that was contrary to the practice of Asian rulers. Indeed Pearson (2003: 114) suggests that it was the Portuguese who effectively introduced politics to the Indian Ocean. Until then, Southeast Asian rulers had not unduly interfered with commerce, as this would have resulted in driving off merchants. The Portuguese used commerce to interfere with religious politics, in particular to disrupt Muslim networks and weaken their power in the Western part of the Indian Ocean; this was looked upon as unethical by local powers. The tributary system that existed between China and some of its vassal states carried the implication of an alliance that included some degree of military protection, and did not have the structure of a monopoly on trade. On the other hand, the Portuguese and later the Dutch also institutionalized the monopoly of goods, a practice that Asian rulers did not support, as rulers were expected to engage in commerce as merchants, not by using political power (section 2.3).

If the Portuguese and the Dutch did manage to have an edge in direct competition with certain Asian powers, this was achieved through a conflation of violence and intolerance. The type of commerce introduced by the Portuguese and later developed by the Dutch, when compared to the pre-existing liberal commerce between Asian powers, has been described as a “vast protection racket” from violence they themselves had introduced (Pearson 2003: 121). Maritime violence was a new element in the region. This is not to say that violence over land and trade had not been present before the arrival of Western fleets. What seems to be the case is that conflicts between Asian powers were largely solved on land, by regular armies, and very seldom were battles fought at sea, with the exception of punitive missions against pirates. Therefore, in particular in

Southeast Asia, fleet wars were largely unknown, since Asian rulers regarded maritime conflicts as a matter for merchants to solve, not a matter of the state.

These unprecedented practices in the history of the region – state-sanctioned violence at sea and monopoly on trade – occasionally afforded Western powers the advantage on which they would build their lucrative networks (Pearson 2003; Newitt 2005).<sup>19</sup> Nonetheless what should be recognized is that Western presence in the Asian region never meant complete domination of local powers by Western colonial structures. This was not true even in the case of India, which was fully exposed to undoubtedly the most organized and powerful Western presence, the British East India Company and the British Empire (see chapter 8).

Compared to the Portuguese, the Dutch had better ships and more efficient navigation techniques, and were better organized. But compared to the fleets of the city-ports, as well as to Chinese ocean-going junks, Western boats must have been rather unimpressive. Through the accounts of early travellers such as Marco Polo and Ibn Battuta, as well as through Portuguese and Dutch sources, we learn that there was hardly any area in which the Western powers were superior to Muslim, Indonesian, or Chinese technology. On the contrary, Muslim and Chinese boats struck the first Europeans as being much larger and more efficient than their own, and had spacious, isolated dry compartments for storing goods, which enabled very long journeys without the necessity of stopping (Levathes 1994). Javanese boats were admired for their lightness and manoeuvrability and the sailors for their skills (Couto 1645). In terms of navigation, there was a deep knowledge of seaways, stellar navigation and trigonometry that Arabs, Chinese, and Southeast Asians seemed to share; the Arabs having been familiar with star navigation from ancient times. Moreover, there is evidence of extremely detailed maps from the Arab world, Malaysia, China, and Korea. By the fifteenth century, East Asia had far more sophisticated maps of the region than early European ones (Ptak 1998). Captain Albuquerque was shown a chart in Javanese script that already traced the Cape of Good Hope, Brazil, Portugal, the Red Sea, and the Persian Gulf, as well as the whole of the Indonesian archipelago and China. A Korean map from 1402, a time when Europeans knew hardly anything of Asia, shows, besides Asia, India, and Africa, a detailed sketch of the Mediterranean and Adriatic seas, with the outline of Italy and the Iberian peninsula (Pearson 2003).

There was, however, one significant difference between the Dutch and Asian fleets, as already noted in the case of the Portuguese: the employment of the Dutch fleets for war. Like the Portuguese, the Dutch had mastered the technique of loading cannons onto vessels and had well-armed crews on board. Until the arrival of these European fleets, city-ports in Southeast Asia were not fortified. Reid (1980) explains this based on an economic rationale. As already mentioned, land was abundant in Southeast Asia, especially the vast

areas of inhabitable coast with easy access to water. The overall population was rather small and life involved constant movement and resettling. Wars, therefore, were usually not fought for territorial conquest but for manpower. As we saw in [section 2.4](#), the most precious capital in the region was measured in terms of slaves, servants, guards, concubines and other people that accompanied wealthy traders and aristocrats. In some notorious cases, these entourages could comprise thousands of people. Therefore, when a city was under attack, if things turned for the worse, the rulers and their entourage would flee, hiding in the jungle until it was safe to return. Cities were really like big villages, with light buildings made of wood; property could be hidden or carried away. The most important point was to hide the people, so that they would not be captured and enslaved by the enemy. If the attacker could not find manpower, he would eventually leave the field. For the same reason, as noted in many early Western and Chinese chronicles, rich merchants and aristocrats were rather unwilling to expose their servants in conflicts, and preferred to hide them rather than risk losing them (Reid 1980: 244–5). As we saw above, this was indeed how Melaka fell to the Portuguese: the sultan and the troops fled, leaving the city to the Portuguese. In short, the Portuguese and the Dutch had a limited technological advantage over Asian powers, most notably in the field of naval warfare (Rosa 2005). In all other respects, they were clearly inferior to their Asian counterparts and surely not in a position to command much prestige.

### 2.5.6 *Final remarks*

Though beyond the scope of this book, the reflections above are interesting for assessing the extent to which local societies might have wanted to ‘model’ themselves along the lines of these early Western ‘barbarians’ (see [chapter 8](#)). If nothing else, they indicate that western countries were clearly not advantaged in any respect, whether social, economic, or technological, in comparison to the Asian powers they encountered in Asia. The Portuguese and the Dutch were in fact rather backward societies, but ones that combined a few dangerous aspects: (i) state-supported religious and racial discrimination which justified armed, violent aggression, and occupation; (ii) limited resources for the population of the time which triggered hunger for new territories; and (iii) a focus on territorial and economic monopoly unfamiliar to local political structures.

We continue the historical analysis with focus on linguistic matters in the [next chapter](#).

## Notes

1 I use autochthonous – rather than anglophone – place names where possible.

- 2 This chapter is meant as a historical orientation for the linguist, in order to provide the context within which speakers and languages function in this region. I have chosen to consult the major scholarly works and historical sources on the topic directly, rather than rely on secondary historical evidence. Nonetheless this chapter is necessarily a distillation of the wealth of socio-historical research produced by historians of the Asian region.
- 3 This should not lead us to postulate an actual segregation between land-bound agriculture and sea-based trade, as the two economic systems operated in fact in a symbiotic relationship.
- 4 Also known as the Moluccas, the Moluccan Islands, the Spice Islands or simply Maluku; the name is believed to derive from the Arab traders' term for the region, *Jazirat al-Muluk* 'the land of many kings' (Ricklefs 1991).
- 5 Temples dedicated to Zheng He can still be found in Palembang, Semarang and Java.
- 6 The prestige of these communities is clear from historical sources. For example, in 1407 Zheng He had established a Cantonese as the head of authorities in Palembang (Kobata and Matsuda 1969); and the greatness of several prominent ports such as Demak, Japara and Gresik – the centre of Islamic culture in the region – is attributed to people of Chinese heritage (Ma Huan 1433; Graaf and Pigeaud 1974).
- 7 Sino-Javanese commerce and Sino-Javanese identity were eventually influenced by a number of factors. Firstly, internal political fragmentation of Java led to and suffered from the rise of a different centre of power, Melaka, as we will see in the next section. Later on, especially during Dutch rule, the Peranakan communities would become once again clearly identified as outsiders, due to the politics of segregation of the Javanese and alliance with the Chinese pursued by the Dutch (Reid 2000). By classifying the Peranakans as Chinese, the Dutch could efficiently exclude Javanese elites from power, and this led to an ethnic duality with dramatic political consequences which remain unsettled to this day.
- 8 The relatively low urban density noted towards the end of the Western colonial era was thus not a reflection of lack of urbanization but rather the result of Western colonial dominance, which led cities under Western control to shrink in population. For example, Melaka under the Portuguese had roughly 12,000 citizens, and only 5,000 under the Dutch.
- 9 This also happened in pre-imperial Europe; city-states such as the Maritime Republics could resort to foreign rulers for international alliances and political impartiality, see section 3.1 this volume.
- 10 Until the arrival of the Portuguese, traders of Chinese origin do not appear to have constituted a numerical majority. Considering the intense contacts between the Indonesian world and China, this is surprising, and it has been suggested that the scarcity of direct accounts of the role of Chinese in local chronicles might well be the result of a negative attitude by Southeast Asian rulers towards the Chinese, as China was perceived as a powerful yet arrogant state (Wang 1996). It may also be the case that it was the power commanded by the Chinese, rather than the numbers, that afforded them the influence they had in the region; this may relate to the extended family structure of their network, often involving many different wives and concubines (see section 2.4). In 1500, the largest group of traders in Melaka consisted of Hindu Tamils, Gujarati, Javanese and Tagalog Muslims (Reid 1993a: 115–6). During the sixteenth century, when Portuguese joined the trade, Chinese and Turkish merchants became more numerous. The Portuguese, and later the

Dutch, were surely influential in raising the profile of non-native groups whom they could use as middlemen to take advantage of local networks, and this particularly favoured the advancement of Chinese merchants in the region (see [section 2.5](#)). It is also likely that Indian and Chinese traders maintained a high status for a longer period of time than other Asian groups, as they were backed by their powerful empires (see Reid 1993a: 118).

- 11 According to Reid (1983), the claim that Southeast Asian slavery was ‘mild’ is the main reason why it remained at the margins of the discourse on slavery until recent times. He argues that, considering the diversity of cultures and beliefs in the region, slavery patterns in Southeast Asia should have been central in that discourse, especially if slavery is treated as a universal phenomenon.
- 12 This is according to *Webster’s Online Dictionary* (2008, accessed April 2008).
- 13 Though forced deportation and sale to third parties were rare until the later days of the city-ports and the arrival of Western powers (Hoadley 1983).
- 14 The first broad class of service distinguished by Hoadley (1983) in indigenous systems of manpower of the region comprises a class of free individuals (peasants) who had the duty to perform a certain amount of labour for their master during their lifetimes (a ‘tax’ or *corvée* duty).
- 15 These were of course not the only Western nations engaged in the East–West trade; all Western European powers, as well as North America, participated in it at some point in time.
- 16 They were said to love their spices more than a jealous lover does his mistress (Sardesai 1997).
- 17 At the same time, historians report how dangerous it was to treat the slaves with undue cruelty, as it was not uncommon for slaves to run amok and turn against their master and kill him (Stavorinus 1798). Note that the word *amok* comes from Malay *amuk* and means ‘crazy’ or ‘furious’.
- 18 In fact, local patterns of bondage were restructured to exclude the more obvious aspects of slavery. Additionally, a major labour force entered the region in the form of immigrant Chinese, and newly formed governments replaced private slavery with different types of state service.

### 3 Linguistic ecologies of Southeast Asia

---

*Language was not a major barrier. Linguistic diversity was part of the everyday experience of Southeast Asian commerce, and virtually all had to resort to some lingua franca.*

Anthony Reid (2000: 159)

As we saw in the previous chapter, the city-ports of Monsoon Asia were characterized by multiculturalism and pluralism, suggesting that they were linguistically diverse environments in which languages such as Javanese, Sundanese, Malay, southern Sinitic (Cantonese and Min), Tamil, and Arabic were dominant in different functional domains. A similar linguistic ecology can still be found in city-states such as Singapore today, as well as many harbours of the region. Inhabitants of these environments are predominantly multilingual, usually raised in households where at least two different languages are spoken, and exposed to other languages from very early on, through the playground, the extended family, domestic servants, etc. In a typical Chinese Singaporean household, for example, parents will most likely communicate in either Singlish, an Asian variety of English with a Sinitic-type grammar, or Mandarin, though parents, and most likely grandparents, may certainly be acquainted with other varieties of Chinese such as Hokkien and Cantonese. There will also be friends who, next to English, speak a South Indian or a Malay variety, and Singlish in fact reflects the presence of these languages in the society, at least in its vocabulary. Nannies are usually involved who speak either a language of Indonesia or the Philippines, or perhaps a South Asian language: their English reflects features of these linguistic backgrounds which also enter the ecology in which speakers interact. There is no reason for us to assume a different scenario in the past. Indeed, quite the contrary is suggested by the fact that older generations in, for instance, Singapore are more multilingual than younger ones (Ansaldi 2004a; Lim 2007). The question is: are such ecologies extraordinary in the history of humankind, so that they produce anomalous social dynamics that may justify special behavioural patterns? And are they only found in the exotic East, and completely different from Western scenarios? The answers, I argue throughout this book, are all clearly in the negative. Markets, whether in terms of ports,



fluvial, or continental settlements, whether periodical or permanent, are one of the defining elements of human culture; they are the loci of not only material but also cultural, sexual, and linguistic exchange.

This chapter introduces the three main players, in terms of linguistic repertoires, of precolonial and colonial Monsoon Asia: Malay, Sinitic, and Portuguese. I will first introduce the evolution of Malay, in particular, the contact varieties of Malay that evolve in trading environments. I will then sketch the specific contexts of the formation of Baba Malay, Cocos Malay, and Singapore Bazaar Malay, articulating the role of Chinese languages in these environments. Finally, I will introduce the Portuguese element, the earliest Western language to enter the region, and the contexts in which Asian-Portuguese vernaculars develop.<sup>1</sup>

### 3.1 Southeast Asia and the role of Malay

Premodern Southeast Asia is an extremely diverse region in terms of languages, with major families all well represented, including Dravidian, Austronesian, Sinitic, Chamic, and Arabic. Within the Austronesian family, the languages of Western Indonesia alone offer great structural diversity and abundant dialectal variation, as well as special-purpose linguistic codes for ritual, literary, and interethnic communication (Fox 2005; Himmelmann 2005).

Today, Indonesians are largely at least bilingual in the standard language (Bahasa Indonesia) and a local language or dialect, though most often they speak two non-standard languages: a language of ethnic identification and a language for interethnic communication. In terms of first or ethnic language, Malay is not and was not the majority language of the Indonesian archipelago. The first place goes to Javanese, with some 37 per cent of all speakers; Malay and Sundanese compete for second place, claiming approximately 13 per cent of speakers each. Large language groups also include Madurese, Minangkabau, Balinese, Buginese, Sasak, and Toba Batak, all of them mutually unintelligible. In Eastern Indonesia alone, one of the most diverse regions, there are some 200 distinct ethnic languages. In total, for a population of a little over 238 million people, there are today 739 different languages, of which two are considered adult second languages (Gordon 2005). There was surely more diversity in the past than there is today, considering that most linguists generally agree that, particularly during the Western colonial era, loss of linguistic diversity was dramatic.

By the 15th century, all the brilliance of this complex archipelagic network ... of diverse urban societies, some Islamic and other Hindu-Buddhist, speaking many languages but all writing a Malay enriched with Sanskrit and Arabic loans as well as Tamil, Mon, Chinese, Persian and Javanese words, was focused on the regional entrepot par excellence, Malacca. (Collins 1996:12)

The quote above is just a reminder of the pluralist multilingual society described in chapter 2 as a prototypical environment in which new languages arose. The most dominant languages of Monsoon Asia used to be Tamil, the language of the South Indian Chola Empire (AD 850–1200), Arabic, the language of literacy, Malay, the lingua franca throughout the region, and Chinese, which was used by official envoys, soldiers, and merchants, and was a politically and socially prestigious language (see [chapter 2](#)). This is evidenced by the existence of a wordlist of 500 Chinese–Malay entries compiled by an archival scribe of imperial China in 1560, with Malay items being largely intelligible to Malay speakers today and not based on literary Malay (Collins 1996: 14–15). This indicates two important things: (a) that Chinese–Malay contacts were ripe in the sixteenth century, and (b) that the language of communication was a colloquial variety based on Malay.

In premodern times, at least two types of lingua franca were spoken, a form of Arabic and a variety of colloquial Malay often referred to as Bazaar Malay (Adelaar and Prentice 1996; Reid 2000). The latter seems to have been the more important and widespread one. We have numerous sources offering detailed accounts of Malay-speaking interpreters who routinely took part in the voyages across the region (Pigafetta 1524; Parmentier 1883). An Italian–Malay list compiled by Antonio Pigafetta in 1522 while travelling through Brunei and Tidore proves that colloquial Malay was in fact used as far as 2,000 kilometres away from the cultural centre of Melaka (Collins 1996: 17). It is for these reasons that the Malay of the time has been compared to Latin in Europe. However, Malay was far more complex in its functional range, as it was the Classical Latin of Southeast Asia, in its written form, used in particular for religious scholarship (and translation), as well as the lingua franca of the Southeast, as a language of trade, international politics, and interethnic communication (comparable to Vulgar Latin according to Collins 1996). Throughout the seventeenth century, Malay was the dominant language of the Malay/Indonesian world and beyond, and it was a requirement for any person engaged in the region to have some understanding of it (Collins 1998: 33). It is obvious that, in its role of lingua franca, it would be difficult to capture this Malay as one single uniform variety; any language of interethnic communication spread over such a vast territory, and spoken by people of so many different linguistic backgrounds, might display a small core of similar features but would otherwise be expected to show a lot of regional variation, as we will see in the historical phenomenon known as Bazaar Malay.

This was a state of affairs not unlike that of *Lingua Franca* in the Mediterranean. Indeed a pluralist and linguistically diverse scenario that is comparable to Asian city-ports can be found in the period of the Maritime Republics that developed in the Italian peninsula from the tenth to the thirteenth century. In Genoa and Venice, Pisa and Amalfi, people of different cultures and linguistic

backgrounds co-existed, i.e. Jews, Turks, and North African traders interacted with Genoese and Venetian merchants, as well as other Northern European parties. Republican settlements with stable populations could be found all over the Mediterranean and as far as the Black Sea. It is most likely in these ecologies that the Mediterranean *Lingua Franca* was born (see Cifoletti 2004; Selbach 2007). The evolution of Malay, and in particular Bazaar Malay, shows interesting parallels both to the Mediterranean *Lingua Franca* and to Vulgar Latin in Europe, as will become clear in the rest of this chapter.

### 3.1.1 *Different types of Malay*

The evolution of Malay varieties is a rather complex matter, partly because reconstruction of older stages of Austronesian languages is far from complete, due to the great diversity of languages in the family (see Adelaar 2005). I am not a trained Austronesianist and am therefore not qualified to arrive at generalizations regarding the evolution of Malay varieties. My aim here is much more modest, viz., to help readers orientate themselves in the complex world of classification of Malay dialects, contact languages, trade languages, etc. This orientation is necessary in order to appreciate the discussion of some Malay contact varieties presented later in this book (chapters 6 and 7). In what follows, I do my best to present the least controversial aspects of classifications and to highlight the areas of agreement among scholars.

The literature often refers to a basic distinction between Malay varieties in terms of ‘high’ and ‘low’. This distinction indicates what appears to be either a diglossic situation between written, formal styles, and spoken/colloquial forms or, in more prescriptive approaches, a distinction between the variety of Malay that should be spoken and the one actually heard in the streets. On closer inspection, however, the dichotomy implied by ‘high’ and ‘low’ appears to fall short of capturing the relationship between the different varieties of Malay that have been in use in the Malay world. Collins (2001) examines the origins of these terms and traces them back to the work of the Dutch minister Francois Valentyn (1726). Valentyn described what he called ‘low’ Malay as the dialect that everyday people understood, while ‘high’ Malay referred to a Malay language that was unintelligible to the ordinary people (Collins 2001: 3). According to Collins, the terminology was derived from the distinction between High and Low in German, associated with the courts of Germany’s highlands and the big cities and trading ports of the lowlands in the north. In later work by Dutch grammarians, more nuanced typologies of Malay varieties emerged, which differentiate between various levels of Malay. What is crucial is that, while in Valentyn’s initial dichotomy, no value judgment was intended, later work by scholars such as Marsden (1812) suffered from the classicist concern with establishing “hierarchies and canons so that an established set

of standard and authoritative principles can be recognized and adhered to" (Collins 2001: 3). With the Romantic tradition came a paradigm shift in which simple dichotomies tended to be rejected, reflected in the work of scholars such as Van der Tuuk, a Melaka-born and Dutch-educated linguist (Collins 2001: 7). He rejected the dichotomy between 'high' and 'low' as too simplistic, claiming that there was no uniform 'low' Malay; he also criticized the 'high' Malay used to translate religious texts by Christian missionaries (van der Tuuk 1856). Instead, he proposed a three-way classification of Malay varieties: (i) centralized, i.e. written language for literary purpose; (ii) local, i.e. regional dialects; and (iii) conventional, i.e. language used for trade. This tripartite division seems more adequate, mainly because it avoids qualitative classifications of these varieties (Collins 2001: 10). What we can learn from these classifications is that more informative distinctions between different varieties of Malay can be found in the social domain, viz., in terms of functions and usage. Indeed, different authors are in general agreement that a distinction between literary and colloquial Malay varieties needs to be made. And, further, that within colloquial varieties one must distinguish between the functions or domains of usage of those varieties that arose from contacts between languages as different as Malay, Sinitic, and European languages ('contact Malay varieties') and those that evolved from less intensive contacts or, perhaps, involving closely related languages, such as vernacular varieties spoken in Malaysia and Indonesia (see Adelaar and Prentice 1996).

One must also be aware of register variation, distinguishing between written and spoken styles. Just like today, a diglossic, or even heteroglossic, situation existed as early as the fifteenth century, in which the colloquial varieties and the written code(s) were far removed from one another (Collins 1996). Valentyn's distinction between a Malay that is understood by locals and one that is not suggests that the different varieties were mutually unintelligible and that the literary language was highly inaccessible to the masses. As a matter of fact, when Muslim missionaries started arriving in the region (thirteenth to fourteenth century AD), they found that a majority of the Malay peoples could not write, which led to the introduction of the Arabic-based Jawi script that replaced the Brahmi-derived Pallawa script (Glynn 1998: 74). A further addition to the difficulty of the written code was that there were several of them available, with many regional varieties and local adaptations (Glynn 1998: 12–13). Perhaps most important is the fact that literature really only developed between the sixteenth and seventeenth centuries; until then it was very limited, especially in the eastern part of the Malay/Indonesian archipelago, where a majority of people did not experience literature in the form of texts but rather in oral form, as songs and poems (Reid 1988: 229). It was only with the growth of the city-ports and with the advent of Islam that written texts became more available. In this respect I believe that one must be careful in

how much emphasis one can put on the role of written/formal varieties in the ecology in which contact Malay varieties evolved. It seems to me that, in the face of historical evidence, we should assume that the features of Malay that were distinctive were selected from the colloquial, spoken varieties rather than from the written codes.

Adelaar and Prentice (1996) distinguish four different periods in the history of Malay: (1) Old Malay; (2) Precolonial Malay; (3) Premodern Malay in the sixteenth to twentieth centuries; and (4) Modern Malay. For the purposes of this study, we are interested in the first three ‘periods’, reviewed below.

### 3.1.2 *Evolution of Malay varieties*

It is generally accepted that a form of Old Malay of the late seventh century AD, probably spoken in southern Sumatra around the Sri Vijaya sphere of influence (see chapter 2), can be considered the ancestor of modern Malay varieties (Adelaar 1994; Adelaar and Prentice 1996). Besides southern Sumatra, other evidence of Old Malay has been found in Java (ninth to tenth century) as well as in the Philippines. This evidence comes in the form of inscriptions in the Pallawa script, a version of the Brahmi script originally from South India (Adelaar and Prentice 1996).

Much later inscriptions (fourteenth century onwards) are usually in the Jawi script, a version of the Arabic script typical of Malay; in this phase we talk about ‘Precolonial Malay’, spanning roughly the fourteenth to sixteenth centuries; this period is also associated with the spread of Islam in non-Malay areas. As already noted, the many loanwords of Malay origin found in languages geographically far apart, such as Tagalog (Philippines), Old Javanese and Malagasy (Madagascar), show that Precolonial Malay was an important language in the region (Adelaar 1994; Adelaar and Prentice 1996). According to Collins (1998: 23), “like Latin in 16th century Europe, Malay was the pre-eminent language of scholarship, commerce, diplomacy and religion”.

In the premodern period (sixteenth to twentieth centuries), Adelaar and Prentice (1996: 674) and Adelaar (2005) distinguish three broad sociolects:

1. Literary Malay: the literary language of texts (in the Jawi script) as well as of aristocratic circles in Malay and non-Malay areas.<sup>2</sup>
2. ‘Inherited’ regional dialects: products of speciation not necessarily involving contact with other languages.
3. ‘Pidgin-derived Malay’ varieties (see [map 3.1](#)).

Note that this periodization is not essentially in conflict with van der Tuuk’s above and is not based on genetic classification but on sociolinguistic observations: moreover, the boundaries between the three are “shaded” (Adelaar 2005: 204). It is the third group that is particularly interesting for my purpose.



Varieties of Pidgin-derived Malay known as 'Bazaar Malay' – a term which reflects the fact that the origins of this variety lie in trade – had existed since precolonial times as languages of interethnic communication. Bazaar Malay is clearly attested in the sixteenth and seventeenth centuries as a language of trade. Even today varieties of Bazaar Malay can be typically heard in their original ecologies, i.e. markets, as can be verified in Singapore, the Riau Islands, Borneo and Malaysia (Collins 2001; Khin Khin Aye 2005; chapter 7 of this volume). Note that the Bazaar Malay varieties spoken today in Singapore and Malaysia are not only languages of the market, but are also used in interethnic communication, among neighbours, and fulfil numerous pragmatic and social functions beyond those usually associated with pidgins (Collins 1984; Khin Khin Aye 2005).

It is with regard to the evolution of Pidgin-derived Malay varieties that we find one crucial difference between scholars. According to some, trade languages such as Bazaar Malay are restructured varieties of Old Malay, most clearly in morphology which is very much reduced possibly due to language contact (see Adelaar and Prentice 1996). Others do not accord an important trigger role to language contact, although they doubt that Bazaar Malay varieties are the outcome of direct internally motivated speciation from Old to more modern varieties of Malay (Collins 1984, 2001). For linguists interested in the evolution of Malay, it is important to appreciate the complex interrelations between different types of Malay.

A particular, subtle differentiation is found in Benjamin (1988), in an attempt to clarify the sociolinguistic variability found in Malay. As has already been mentioned, Old Malay is most often attributed to the Sri Vijaya Empire (seventh to fourteenth century AD). In this period, a number of Malay-speaking polities developed, characterized by aristocratic courts surrounded by tribal populations (Benjamin 1988: 13). In these settings we find a cline of Malay varieties from the most cultivated to the most tribal. Crucially, the latter are, according to Benjamin, not 'low' varieties but rather the closest ones to Old Malay. This is because the courts were more often than not established with significant foreign elements. Indeed colloquial Malay varieties and Old Malay share significant grammatical features in common, while the Classical Malay developed in the precolonial period can be seen as a constructed language for the purpose of representing the courts' legitimacy to the outside (Benjamin 1988: 16). The main structural differences are the following: colloquial Malay does not overtly express thematic roles, agency, and transitivity, resulting in high frequency of naked roots; paratactic structures are frequent, and the discourse has an egocentric dimension, following topic–comment patterns.

This does not mean that it is a simplified form of literary Malay varieties, where thematic roles, agency, and transitivity are overtly marked. On the contrary, Benjamin argues that colloquial Malay is far more 'condensed', requiring



the speaker to construct a more sophisticated and diverse range of interpretations. What is crucial is that these differences are motivated by the immediate ecologies in which they evolved. Colloquial Malay varieties are associated with mutual cooperation in communication and sharing of special knowledge, and function within a context that offers cues for correct interpretation. Classical Malay on the other hand evolved for articulated expression between distant parties and allows specification of social relations where contextual information is absent. Benjamin's interpretation suggests that the relationship between 'high' and 'low' varieties may simply be a diglossic one, and implies that colloquial Malay is not necessarily a language evolved through simplification of literary varieties (see also chapters 7, 8 and 9).

### 3.2 Malay contact varieties<sup>3</sup>

Whether a single well-defined pidgin Malay ever existed is doubtful, as already expressed in Reinecke *et al.* (1975). This is still true of Bazaar Malay varieties, which are widespread in the region, although one may claim that they are contemporary manifestations of the early Malay lingua franca. While some Bazaar Malay varieties are still spoken predominantly at markets of the region, other Bazaar Malay varieties have become indigenous languages of specific communities (Adelaar 2005). For example, Baba Malay, the language of mixed Chinese–Malay communities (see [section 3.3](#)) might be seen as a nativized variety of Bazaar Malay (Reinecke *et al.* 1975). In fact Shellabear (1913) had already proposed that Bazaar Malay and Baba Malay were essentially the same language, and ventured that the varieties known as 'restructured' Malay were created by the Chinese in Malaysia. In the most comprehensive study of Baba Malay to date, Lim (1988: 5) suggests a close relationship between Baba Malay and Bazaar Malay. As I will show in chapter 7, the similarities between Baba Malay and the Singapore variety of Bazaar Malay are indeed quite strong.

However, the historical phenomenon of Bazaar Malay (sometimes also referred to as 'Trade Malay') cannot be directly equated with any existing variety, because, as far as I know, it was never the language of one specific community geographically and culturally defined, such as the Babas or the Cocos Malays (see [section 3.3](#) below).<sup>4</sup> I will argue that it is better to think of Bazaar Malay varieties as sharing a set of communicative strategies used between different communities, or else used by the very large and diffuse community that traded along the Malay/Indonesian maritime routes, a different kind of social network than the one associated with an ethnic group.

When looking at contemporary Bazaar Malay varieties, it is very useful to state explicitly the ecologies of specific varieties. For example, Collins (1984: 164) argues that the Bazaar Malay in use in Malaysia fulfils a number of diverse functional needs beyond basic communication between speakers of



mutually unintelligible languages, and shows interesting syntactic and phonological innovation. Similarly, Khin Khin Aye (2005) finds that Bazaar Malay as spoken in Singapore has a wide range of functional uses and fully allows for sophisticated types of communication. As already mentioned, in comparing Baba Malay to Bazaar Malay in Melaka, Lim (1988) finds many grammatical similarities between the two and proposes that we differentiate between the two varieties as follows: Bazaar Malay is used by the non-English-educated as a language of interethnic communication; Baba Malay is used today typically by well-educated communities of a homogeneous ethnic origin. While Baba Malay has an extensive Hokkien lexicon that clearly identifies one of the input languages (chapter 7 of this volume), Bazaar Malay has lexical influences from many languages (Khin Khin Aye 2005). Lim (1988) also notes that Baba Malay is more focussed as a linguistic system, while Bazaar Malay shows a lot of individual variation. In other words, Baba Malay is an in-group vernacular, while Bazaar Malay varieties are typically used as lingua franca, for interethnic communication. The fact that Baba Malay is readily identifiable as a single, discrete linguistic system is thus a consequence of its ecology: it has the features that set it apart because it has evolved in a community that, through admixture of foreign (Hokkien) traders and indigenous people, has come to form a discrete cultural entity (section 3.3 below and chapter 7). This condition is not necessarily fulfilled in other market settings, where different groups in contact influence Malay in different ways resulting in less focussed situations.

The hypothesis I am advancing here, elaborated in chapter 7, is thus that linguistic speciation is fundamentally the result of the particular evolutionary patterns that occur in different ecologies (Mufwene 2001). When discussing Bazaar Malay, we should be careful to recognize that there are really two different linguistic phenomena here: (1) the historically attested diffuse Bazaar Malay lingua franca spoken across the Malay/Indonesian archipelago as a code of trade and interethnic communication; and (2) specific varieties of Bazaar Malay spoken today in different cities across Southeast Asia, which show different degrees of focussing and different typological traits depending on the speech community.

In constructing a structural profile of the Bazaar Malay lingua franca, it is useful to consider the notion of Pidgin-derived Malay (PDM) developed by Adelaar (2005) to refer to what appear to be local Malay vernaculars and restructured varieties such as Baba Malay, Chitty Malay (Malay–Indian mixed language), Java Malay, and East Indonesian Creoles.<sup>5</sup> Note that the notion of PDM does not imply any firm claims as to the existence of an actual Malay pidgin as a precursor of other Malay contact languages. The notion of PDM must be interpreted in the sense of PDM *varieties*, in line with the diffuse nature of the lingua franca mentioned above; these varieties are related by a cluster of mostly grammatical features that seem to be shared by a number

of contact varieties across a vast geographical region and by many different ethnolinguistic groups. These include: possessive constructions based on *punya* ('owner' > 'to possess') [possessor + *punya* + possessee]; very few productive morphological affixes compared to Literary Malay (i.e. isolating typology); pronouns of Sinitic, Hokkien origin, especially first and second person; plural pronouns formed with singular form + Malay *orang* 'person'; reduced forms of demonstratives that precede the noun: *ini* 'this' and *itu* 'that'; use of the Malay existential marker *ada* to indicate progressive aspect; 'give' and 'make' causative constructions, based on Malay *kasi* 'to give' and *buat* 'to make' preceding the verb; polyfunctional preposition *sama* (these are analyzed in chapter 7).

The notion of PDM moves us away from the danger of 'constructing' a single linguistic code that perhaps never existed, as might also be the case for Lingua Francas of the Mediterranean (Selbach 2007). Instead, PDM suggests the existence of a variety of similar linguistic codes which arose for the purpose of commerce between speakers of different languages. When we look at the accounts and analyses available on Bazaar Malay, this picture is confirmed by the variation found between speakers of different ethnic groups, as noted in Collins (1984: 165), Lim (1988: 50–2) and confirmed by Adelaar and Prentice in their survey of PDM (1996: 675–6).

PDM varieties constituted an important cultural element of the region, serving as varieties for interethnic communication throughout Monsoon Asia (see chapters 6 and 7). It is clear that PDM features constitute an important ingredient of many contact languages of the region that evolved in precolonial and colonial times, as seen in Sri Lanka Malay, Cocos Malay and Baba Malay, as well as Singlish (chapters 6 and 7). These features are likely to be of ancient origin, as predicted by the founder principle, according to which, under particular evolutionary circumstances, older features gain in robustness and tend to be maintained in the histories of some languages (Mufwene 1996b). Considering the Sinitic influence in some PDM features (see chapter 7), this corroborates the idea that a degree of Malay–Sinitic contact was involved in the early evolution of PDM varieties. In this sense, the notion of PDM is useful because, by not assuming the existence of one unitary code as language of trade and interethnic communication, it allows us to consider a more sophisticated and realistic option. PDM varieties may be regarded as the outcome of various processes of language contact between different populations.

The range of Malay contact languages, trade jargons, and pidgins that are known to have existed to date, as well as their geographical spread, indicates without a doubt the significant role that Malay must have had in the region. For a region that, in precolonial times, had a very small population (Reid 1980), the number of mixed varieties found also attests to the intensity and geolinguistic diversity of contacts between different populations. Moreover, the number of different Malay registers still found today reveals the complex sociolinguistic

functions it served. In what follows I introduce some of the more interesting and better-known varieties discussed or referred to in this book. I limit myself to the Strait of Melaka and the island of Java, which are the most interesting regions for this history; for a detailed and complete list, see Adelaar and Prentice (1996: 676–93).

### 3.2.1 *Contact vernaculars*

- i. Chitty Malay: spoken by mixed South Indian–Malay descendants in Malaysia.
- ii. Baba Malay: spoken by the Peranakan communities of mixed Chinese–Malay/Indonesian descent in Malaysia (see chapter 6).
- iii. Jakarta (also Batavian or Betawi) Malay: spoken between different Indonesian groups, between Chinese and Indonesians, between masters and slaves, among Chinese belonging to different Sinitic groups, between Chinese and non-Chinese, and also by Eurasians when shifting from Portuguese to Malay (Grijns 1991). This was possibly already the mother tongue of Indonesian Peranakans. It may have been in existence since the eighteenth century (Collins 1995; Nothofer 1995) and is spoken today by 1 million people.

Baba Malay will be described in detail in section 3.3 and chapter 7 of this volume, as a prototypical example of a contact language of the region.

### 3.2.2 *Lingua francas*

- iv. Bazaar Malay: market language usually spoken by older generations as language of interethnic communication. Ubiquitous to urban centres across the region, in Singapore it is being replaced by Singapore English in the new generations (see chapter 7 and map 3.1).
- v. Sabah Pidgin Malay: used between Chinese and Sabahans as well as between Sabahans of different ethnolinguistic background in Malaysia.
- vi. Tangsi Malay: used to be spoken by the Dutch East Indian army and contained Dutch, Malay, and Ambonese elements. It was exported to the Netherlands as Melayu Sini.
- vii. Java Malay: still widespread in Java, used to be the language of lower administration (see also Cocos Malay, chapter 6).
- viii. Java Port Malay: trade language of the ports on the northern coast of Java. It is extinct and no records are available; it may have been the ancestor of Jakarta Malay.

The differences between these varieties are often most marked at the phonological level, and peculiar regional grammatical influences can at times also be identified (Adelaar and Prentice 1996: 675). The existence of many different

forms of PDM across the region indicates that these varieties were used by all ethnic groups, whether Malay, Indian, or Chinese. In fact, the influence of Malay went as far east as Papua New Guinea, where it became the lingua franca of Chinese and Javanese workers on several plantations in the nineteenth century. There are also Malay loanwords in Hokkien and in Taiwanese texts from the seventeenth century (Adelaar 1994; Adelaar and Prentice 1996). Its presence is attested also to the south, on the Cocos (Keeling) Islands and Christmas Island off the western coast of Australia (see chapter 7). To the west, features of PDM are identifiable in Sri Lanka Malay (chapter 5), in Malagasy, and in South Africa, where Cape Malay is claimed to have influenced Afrikaans (Adelaar and Prentice 1996).

It is difficult to draw a clear distinction between vernaculars and lingua franca. As has often been remarked upon, Baba Malay and Bazaar Malay varieties share many common features (see the comparison with Singapore Bazaar Malay in chapter 7, section 3). This is not surprising: for one thing they are geographically very close; more importantly, Baba Malay is closely associated to the history of Melaka and of the Peranakan communities found in Singapore and Penang (Malaysia). Melaka is also historically well positioned for being one of the cradles of Bazaar Malay varieties and cross-fertilization between Bazaar Malay and Baba Malay surely occurred. As merchants and mediators between Chinese interests and regional powers in precolonial times, and later as middlemen between those same regional powers and colonial interests, the Babas were an influential element in the ecology of Bazaar Malay. Baba Malay and Bazaar Malay also show closest adherence to PDM features; in itself this suggests that PDM features could be seen as the core (more stable and widespread) properties of Bazaar Malay varieties across the region, found as far east as Ambon (van Minde 1997) and as far south as the Cocos Islands (see 3.3.2). As the grammatical comparison in chapter 7 will show, Baba Malay and Bazaar Malay as spoken in Singapore can only be clearly differentiated in terms of lexicon and phonology. Baba Malay has a clearer Hokkien imprint and is phonologically more convergent with Malay, while Singapore Bazaar Malay has more variation, being less focussed. In fact there is variation across different ethnic groups in (Singapore) Bazaar Malay, not only in terms of phonology (Lim 1988) but also in terms of syntax, as shown for the case of Indian Singapore Bazaar Malay in Sasi Rekha (2007).

### 3.2.3 *Registers*

- ix. Pesisir ‘coastal’ Malay: developed in the sixteenth century as a literary form for Muslim instruction.
- x. Chinese Malay: literary style based on spoken Malay (and Java Malay) developed by journalists in the cities of the Dutch East Indies in the nineteenth century.

- xi. Urban or colloquial Malay: informal, spoken language (as opposed to literary Malay). This is the *de facto* Malay language of everyday usage.

If we were to include the island of Sumatra to the west and the islands to the east of Java (Sulawesi, the Maluku Islands, etc.), this list would obviously increase. The proliferation of so many different mixed varieties is a result of what was, and still is, a linguistically heterogeneous environment in which multilingualism has always been the norm. As we will see in [section 3.3](#) below, as well as in chapters 6 and 7, while Malay was undoubtedly a major influence on other languages of the region, it was also influenced by some of these. In particular, I will focus on the role of Sinitic in the evolution of some mixed varieties of the Malay/Indonesian region. For the time being, it is particularly important to remember that, when talking about Malay influence in language evolution of this region, we are not referring to Literary Malay, but rather to one of the many colloquial, probably mixed varieties spoken in Malaysia or, more likely, Indonesia, from around the fifteenth century onwards (see chapters 5, 6 and 9).

### 3.3 Introducing contexts of formation

This section sketches three specific contexts of CLF in Monsoon Asia; the evolution of these varieties and the wider ecological significance that emerges from these contexts will be discussed in the following chapters, in particular chapters 6 to 8.

#### 3.3.1 *Baba Malay*

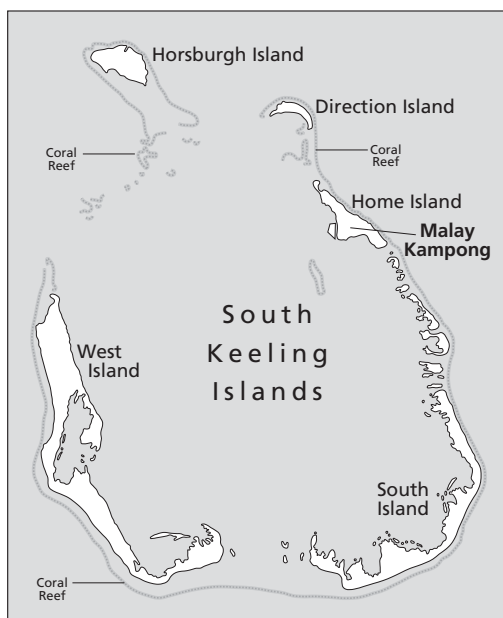
Since the general historical settings of city-ports such as Melaka have already been discussed in chapter 2, I focus here on the social networks within which Baba Malay evolved. As we saw in chapter 2, the contacts between Malay and Indonesian people and Chinese merchants go back to at least the early fifteenth century in the history of Southeast Asia. As explained in section 2.2, the fifteenth century was merely a turning point in the expansion of Chinese overseas communities. Thus, Chinese languages had been present in the region since before the arrival of Western powers, when the massive fleets of Zheng He visited Java and Melaka, leading to an increase in commerce and the local Chinese population and giving rise to permanent Chinese settlements along the northern coast of Java. The Chinese maritime commerce in Southeast Asia continued to develop in the sixteenth and seventeenth centuries, led by the Fujian-based traders Zheng Zhilong and Koxinga, among others (Pan 2000: 49).

As we saw in chapter 2, in the city-ports, Chinese males frequently formed unions with Malay and Indonesian women, slaves, and concubines (see Lim 1988: 9; Tan 1988a, b; section 2.4 of this volume). These early unions were an

important factor in the acculturation of the Chinese and the subsequent emergence of the Peranakan community and the Baba Malay language and culture (Tan 1988a: 42f.). The fact that the Chinese married non-Muslim women, often slaves of Balinese, Bugis, or Javanese origin, meant that no conversion was required, and this led to maintenance of some Chinese traditions (Skinner 1960; Tan 1988a: 40–1; Pan 2000; Ansaldo, Lim and Mufwene 2007). The Malay/Indonesian cultural element of Baba communities is, however, very clear, as evidenced, for instance, by their dress and cuisine (chapter 7). While the origins of these communities lie in a practice of interethnic relations, their later evolution is characterized by an increasing endogamous practice. In these mixed, Peranakan communities, the Nyonyas ('female members') were encouraged to marry Babas ('male members') (Lim 1917: 876) and, by the time the communities had reached a critical mass and the sexes had equalized (nineteenth century), unions between the Babas and the local women also decreased. In fact the communities became less receptive to non-Chinese, and marriages occurred predominantly between Peranakans or between them and new immigrants from China (Lim 1917: 867). Their language, Baba Malay, combines Malay and Hokkien features (Lim 1988; Ansaldo and Matthews 1999; Ansaldo, Lim and Mufwene 2007). The Babas of Melaka and Singapore (as well as Penang) are thus clear examples of innovative cultural admixture in Southeast Asia. In all aspects of tradition, be they artistic, literary, religious, or linguistic, the Peranakans developed unique traits by combining elements of Malay and Chinese cultures that set them apart from other Chinese, the more indigenous local populations, and other ethnically mixed groups (Tan 1988b; Rudolph 1998).

### 3.3.2 *Cocos Malay*

The Cocos (Keeling) Islands consist of a coral atoll of twenty-seven islands located in the middle of the Indian Ocean a thousand kilometres southwest of Java (see [map 3.2](#)). The origins of the community are particularly intriguing: the first to settle on the atoll was Alexander Hare, who had been working with the sultan of Banjarmasin in Borneo until he ran into trouble either with the sultan himself or with the East India Company. From Borneo he fled, among other places, to South Africa, with a stopover in Melaka where he acquired part or most of his followers. The women in the group that settled the atoll in 1825 were virtually Hare's *harem* while the men acted as servants and labourers. In 1827 John Clunies-Ross landed on another island of the atoll, accompanied by his family and a crew consisting of mostly Malay and Indonesian people; he apparently disapproved of Hare's lifestyle (Gibson-Hill 1947: 149) and eventually proved to be a more reliable 'master' for many of the Malay/Indonesians in Hare's company. In 1829 the original population totalled 98



Map 3.2 Cocos (Keeling) Islands

people mostly of disparate Malay or Indonesian origins, including women from Sumatra, Borneo, Java, Bali, Timor, and New Guinea. There were also South Africans, Chinese, and Indians and a few British people. They were portrayed as speaking a form of Malay similar to the dialect of Batavia but seemed to lack knowledge of a number of words for which they had invented their own by modifying English or Dutch words (Gibson-Hill 1947: 168). After Hare's death, Clunies-Ross and his heirs acted as benevolent rulers of this community until the 1950s, when the islands became Australian territory. His eldest son married a local Malay and was succeeded in turn by his eldest offspring. Charles Darwin visited the atoll in 1836 for eleven days (Gibson-Hill 1947: 150) and noted that:

[the] Malays are now nominally in a state of freedom and certainly are so as far as regards their personal treatment; but in most other points they are considered as slaves. From their discontented state, from their repeated removal from islet to islet, and perhaps also from a little mismanagement, things are not very prosperous. (Darwin 1842: 542)

He also records that "[t]he natives belong to different islands in the East Indian archipelago, but all speak the same language ..." (1842: 548). The naturalist Forbes (1885) noted the mixed physical appearances of many of the natives,

which showed at times Asia, at times Papuan, at times African provenance. This is not an exaggeration: families with clearly diverse racial roots are still very much evident in the community today and readily acknowledge their distant heritage. Despite this heterogeneity of their ancestries they identify themselves as *orang java*, Javanese people.

In the nineteenth century, in order to develop the production of coconut oil, labourers from Eastern Java were imported, followed by contract labourers (ex-convicts) from Batavia. This led to the establishment of a second settlement or *kampong*, some distance away from the Cocos compound. By the mid-nineteenth century, these people of Javanese origins constituted half of the population. Throughout this period a form of Malay had been in use between 'masters' and labourers since both John Clunies-Ross and Alexander Hare were said to be fluent in Malay (Gibson-Hill 1947; Cocos community p.c. June 2002). Their association with Melaka and various sultanates of Borneo suggests that, as traders and most often in contact with local populations, sailors, and slaves, they would have had knowledge of Bazaar Malay varieties, and that Dutch and English, as well as languages of Indonesia, would have been represented. Apart from a migration of roughly 800 individuals to West Borneo and Western Australia in the twentieth century (due to overpopulation of the small island), the community stayed rather isolated from the rest of the world until 1955, when it became first part of the Straits territories, together with Singapore, and when Malay teachers visited the islands. In the 1980s, it became part of Australia, a strategic outpost in the Indian Ocean before the era of satellite technology. In the 1970s, education was introduced, predominantly in English but with bilingual provisions, in the sense that Indonesian was offered as a second language. In more recent years, it has been debated whether to replace Bahasa Indonesia with Bahasa Melayu.

From the description above, it can be seen that the ecology in which Cocos Malay evolved was for the longest time rather stable and closed to the outside world, with linguistic features of diverse origins available to speakers. Over the past three to five decades, exposure to the rest of the world has occurred, and Standard Malay (Bahasa Melayu) and Standard Indonesian (Bahasa Indonesia) as well as Australian English have entered the ecology, not only through population movement and education but also through satellite TV, popular culture, music, etc. The current population of the Cocos-Keeling consists of approximately 600 individuals residing on the two inhabited islands of the atoll. West Island is home to 100 Australians, mostly administrators and schoolteachers on two- to three-year postings. Home Island, the original settlement of the first settlers, still houses 80 per cent of the population. These are the Cocos Malays, descendants of the labourers brought to the islands during the 150-year period in which the islands were an estate of the Clunies-Ross clan (British East India Company). The community is Muslim and speaks 'Cocos Malay', a PDM



according to Adelaar and Prentice (1996) with strong Javanese influences, as well as Australian English, especially in the younger generations. They refer to their own language simply as 'Java' or 'kampong Java' (Javanese of the community), a practice also found in the Malay Diaspora of Sri Lanka (see chapter 6).

While in general the Cocos Malays exhibit a mixed Muslim–Malay culture, a peculiar aspect of their material culture lies in the area of boat building. The type of boat they build, called *jukung*, is an interesting adaptation of British (Scottish) and Malay features (Bunce 1988) to the lagoon environment in which it is used. It has nearly vertical sides, a flat bottom and miniature-sized keel, and bow and stern are almost vertical so that the boat can be propelled forward and backwards with equal ease. These features are perfectly adjusted for navigation in the shallow, choppy waters of the internal lagoon. The steering is like the one typical of English boats with rudder and lines. While the boats have one mast and two sails which can be governed by a single man who can also fish at the same time, because of the short keel, children are usually used onboard as ballast. It is clear that this boat evolved as a combination of the nautical expertise of the Clunies-Ross clan and the workmanship of the Cocos inhabitants applied to local conditions. In 2002, there were only few individuals left capable of building a traditional *jukung*. The hybrid nature of these boats, in terms of architecture and materials, is an aspect of cultural fusion typical of new communities in which new languages emerge (see chapter 7).

The basic typological aspects of Cocos Malay and the relation to other contact Malay varieties are illustrated in chapter 7.

### 3.3.3 *Singapore Bazaar Malay*

The linguistic ecology of Singapore is rather well known to the world of contact linguistics due to the amount of investigation and reconstruction it has received (see in particular Gupta 1994; Lim 2004, 2007). What follows is thus a brief overview of the formative years of the city-state, as well as some reflections on its current linguistic ecology.

The island of Singapore was acquired by the British East India Company in 1819 as part of the Straits Settlements, which included Penang and Melaka. As a strategic location for the China–Southeast Asia trade, it already had a heterogeneous population, which, by 1820, consisted of over 5,000 inhabitants. These included Malaysian Malays, Javanese, Balinese, and Buginese from Indonesia, Indian Tamils, Malayalees, Punjabis, and Sikhs, as well as Hokkiens, Teochews, Cantonese, Hakkas, and Hainanese from China; mixed groups such as the Peranakans and the Eurasians were also represented (Lim and Foley 2004; Lim 2007, forthcoming). While initially the largest proportion of the population were 'Malays' (from various parts of Malaysia and Indonesia) followed by the

Table 3.1. *Linguistic ages in Singapore (adapted from Lim 2007)*

| Age                                                 | Socio-history                                                                                                                              | Varieties <sup>6</sup>                                                      |
|-----------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------|
| I. Immigrant substrates<br>Pre-1800s – mid-1970s    | Multilingual, heterogeneous society Bazaar Malay spoken by all groups Hokkien                                                              | Bazaar Malay and Hokkien<br>Sinitic and Malay varieties<br>Indian languages |
| II. Official languages<br>Mid-1970s – late 1980s    | Shift towards fewer official languages. Bazaar Malay and Hokkien decline, replaced by English. Mandarin displaces other Sinitic varieties. | English<br>Mandarin<br>Hokkien and Cantonese<br>Bazaar Malay Tamil          |
| III. Global media languages<br>Late 1980s – present | Varieties of English develop<br>Cantonese influence increases<br>Mandarin-only attitude relaxed                                            | Englishes<br>Mandarin<br>Cantonese<br>Other Sinitic                         |

Chinese traders, over time the Chinese communities grew rapidly in number to comprise half of Singapore's population by the beginning of the twentieth century (Lim 2007). Throughout this entire period, i.e. the nineteenth century until around the 1970s, a form of Bazaar Malay was used as the primary interethnic lingua franca, with Hokkien the main intra-ethnic lingua franca for the Chinese (Lim 2007). Over time, English grew as a medium of instruction, gradually replacing instruction in other languages (in particular Malay and Mandarin) and established itself as a new language of interethnic communication. This led to the evolution of another contact phenomenon in Singapore: Singapore colloquial English (Platt and Weber 1980; Gupta 1994; Lim 2004), also known as Singlish, a variety in which Sinitic, Malay, and English grammars blend to form what is *de facto* one of the two languages most Singaporeans today grow up speaking; the other being Mandarin (Ansaldi 2004b; see also Bao 2001, 2005; Bao and Lye 2005). A number of other languages are also currently in use in Singaporean society: besides various registers of Mandarin and English, Hokkien is still spoken, though it has less prestige than it once had when it was the language of the dominant Chinese community. Cantonese is well represented and very vital, as argued in Lim (2007, forthcoming). Malay is still the official language although it is spoken by relatively few people outside the Malay community as a primary language, but is still very much in use as a street language – i.e. Bazaar Malay – in markets, with vendors, and in food courts; Tamil is used within the Indian population. Lim (2007) captures the linguistic history of Singapore in terms of three historical stages or 'ages', synthesized in [table 3.1](#).

For the purposes of this section, it is noteworthy that Singapore's ecology, though still very multilingual, has changed over time from being highly multilingual at the societal and individual levels, with a strong dominance of Malay

and Sinitic features, to largely English–Sinitic, with a decrease in the usage of Bazaar Malay. This can be seen as a general trend in the postcolonial history of the region, as a consequence of the introduction of standardized and global languages. However, as documented in Khin Khin Aye (2005: 24), Singapore Bazaar Malay is still alive, mostly used by elderly Singaporeans of Malay, Baba, Indian or other non-Malay origins, when other common languages are not available, while in the younger generations of Singaporeans Singlish functions as a language of interethnic communication. Singapore Bazaar Malay is thus not only used in markets and shops but also in everyday conversations between neighbours and in interactions with domestic helpers (Khin Khin Aye 2005: 25). The structural features of Singapore Bazaar Malay are discussed and compared to Bazaar Malay and Cocos Malay in chapter 7.

### 3.3.4 *Thoughts on evolution in the Malay world*

The three contexts presented above illustrate the fact that Bazaar Malay varieties – or PDM varieties – were essential languages of interethnic communication used within Malay and Indonesian communities, as well as between these and other groups, in particular Chinese and Western traders. At the same time, it is important to note the significant differences found in the three contexts of formation: Baba Malay emerged as part of a process of new identity formation in which a new cultural identity was defined; Cocos Malay evolved in relative isolation from the other languages; and Singapore Bazaar Malay exists as a diffuse code with specific social functions. As I will show in chapter 7, these different ecologies have had an impact on the linguistic system that has evolved in each of them.

Although PDM features are a useful historical diagnostic for understanding evolution in the Malay world, the notion of ‘pidgin’ carries with it a number of problematic assumptions. For the sake of simplification, there are two typical ideas associated with pidgins and pidginization: (a) it involves economizing tendencies that occur in second language acquisition and/or in informal contexts with a good deal of shared knowledge; (b) it is a necessary requirement for the emergence of a new language (the pidgin-to-creole scenario). While (a) can be useful in understanding aspects of contact-induced change, as long as the ecology of the contact environment is taken into consideration (but see chapters 7 to 9), (b) implies the problematic idea of a pidgin becoming the native language of a new generation and thus turning into a Creole. Crucially, this idea, propagated in particular by Bickerton, has been not only difficult to prove in the history of language genesis but also proven fundamentally wrong for the evolution of Hawai‘ian Creole English (Roberts 1998, 1999, 2000; Siegel 2000, 2007).

As already noted by Le Page and Tabouret-Keller (1985: 190), nobody grows up in a ‘pidgin-only-speaking’ society. The multilingual ecology in

which contact Malay varieties evolved, characterized by widespread bi- or multilingualism, clearly suggests that restricted codes for the purpose of trade and fully fledged languages would have existed side-by-side (and furthermore overlapped in a number of social domains). Therefore a pidgin-to-creole evolutionary scenario is ruled out in the evolution of PDM varieties. While the idea of PDM features rightly captures the nature of what would have been a diffuse and widespread linguistic code for the purposes of interethnic communication, we should be careful about the use of 'pidgin' and therefore avoid theoretical and ideological biases.<sup>7</sup> As argued in Ansaldo, Lim and Mufwene (2007), for example, there is nothing exceptional about the evolution of Baba Malay in Peranakan communities: there was no break in transmission, no lack of fully fledged target languages, no unequal social relations. It is a clear example of the fact that contact language formation is an aspect of the creation of a new cultural system (Chaudenson 2001).

Similarly, pidginization as structural simplification is difficult to invoke in the evolution of Baba Malay and Bazaar Malay, since the typological congruence with isolating Sinitic typology and the general absence of morphologically rich literary Malay in the contact setting fully account for the structural output of the contact language formation process. Also, the use of the term *pidgin* could be used to stigmatize varieties associated with such an evolution in a number of ways: as genetically impure or orphan languages, as simple or insufficient for proper communication and therefore worth abandoning, and as outsider languages within an ethnic discourse of purity.<sup>8</sup> We will see that none of these aspects need be ascribed to languages such as Baba Malay, Cocos Malay and Makista.<sup>9</sup>

As will become clear in chapter 7, and as argued in Ansaldo (2003b), Cocos Malay is clearly less Sinitic than Baba Malay and Bazaar Malay. This simple observation happens to shed significant light on the evolution of PDM varieties. The origins of the Cocos people go back to the early eighteenth century. The language spoken in the early years of the community was a form of Bazaar Malay, possibly from Melaka. Later in the nineteenth century, Java Malay came into the picture. There are two options to account for the difference between, on the one hand, Baba Malay and Bazaar Malay and, on the other, Cocos Malay: (a) Cocos Malay had a core of Bazaar Malay features (the PDM traits) to which later Malay morphological processes from Java Malay were added. The core features are (in part) Sinitic-derived, the Malay morphology being acquired later; or (b) Cocos Malay was from the beginning the way it is today – Java Malay only reinforced its features. Therefore the Bazaar Malay variety spoken by the Cocos people was more Malayic than Singapore Bazaar Malay and Baba Malay.

What is crucial here is that neither scenario implies a pidginization process from Literary Malay to a colloquial Malay variety along the lines of morphological reduction. If Cocos Malay acquired Malay morphology through contact with Java Malay, this is a process of morphological development, usually not

observed in pidginization. If Cocos Malay was the way it was from the beginning, we may be looking at one of the best preserved examples of an early colonial Bazaar Malay variety. Recall the geographical isolation under which the islanders lived until the 1960s, with the exception of contact with Javanese. Considering that a majority of the morphological material described is present in the oldest generation, it is quite likely that most of it is not of recent development but would date back to either the Bantamese labourers or to the original population. In the latter case, we would be looking at a situation of prolonged maintenance of a combination of PDM features and aspects of Malay morphology, a hybrid system resulting from the contact between Bazaar Malay features and Malay and Indonesian vernacular varieties.

### **3.4 The role of Portuguese in Southeast Asia and Southern China**

Bazaar Malay pre-dates the arrival of Western powers and, in its various forms, it has continued to play a relevant role in the ecology of the region, as seen in Sri Lanka Malay (chapter 6) and the Malay–Sinitic varieties. However, we need to acknowledge that Portuguese also had quite a significant role in this ecology, as evidenced by the existence of Indo-Portuguese and Malayo-Portuguese varieties. First of all, Portuguese was the *lingua franca* among Western traders in Monsoon Asia (Pearson 2003: 156–7). Secondly, Portugal was the first Western power to develop significant trading networks in the region, a fact that left a long lasting linguistic influence in various parts of Asia, notably India, Sri Lanka, Melaka, and Macau. It was thus the first Western language that Asian populations came into contact with (Clements 2009). Further, as already outlined in chapter 2, because of a basic deficiency in manpower, the Portuguese ‘empire’ remained heavily localized, decentralized, and deeply integrated, out of necessity, with local political structures. Portuguese settlements were thus characterized by small numbers of Portuguese individuals and a high proportion of local women, servants, and soldiers, a situation that encouraged intermarriages.

The effects of Portuguese presence can be seen in the linguistic ecology of the region: Portuguese-based contact varieties by far outnumber the contact languages in which either Dutch or English played a role; the Portuguese language had a wider and deeper influence in the ecology of Asia also thanks to the fact that it was used as a *lingua franca* by other Western traders.<sup>10</sup> As argued in Bolton (2003), Portuguese influence in fact carries all the way into English – it is through Asian Portuguese that many new lexical items came into the young, unstable, highly variable English of the time. This must be kept in mind when we look at the later influence that English had in Asia. The Asian Englishes that emerged in places like Singapore are not merely a modern development due to the impact of English on local languages; they may perhaps be viewed as a continuation of mutual linguistic exchanges occurring

in East–West encounters from the seventeenth century onwards (Bolton 2003: 122–96; Ansaldo, Lim and Mufwene 2007). Portuguese therefore occupied an important role in East–West exchanges from India to the South China Coast, where we know that knowledge of Portuguese was essential for non-Portuguese Western traders in their dealings with Chinese authorities (see chapter 8). Likewise, in Indonesia the Dutch relied heavily on Portuguese along with Bazaar Malay, which explains why Dutch never featured prominently in the ecology of Southeast Asia (Taylor 1983; chapter 2 of this volume). What is noteworthy is that ‘Portuguese’ in the Asian region really refers to Asian Portuguese (Indo- and Malayo-Portuguese), as we will see in the case of Makista, the vernacular Portuguese variety of Macau.

As I explain in chapters 7 and 9, it is important to reflect on the type of Portuguese that was brought to Asia; this is part of a more general, methodological concern with using historically plausible varieties in comparative linguistics addressed throughout this book. Le Page and Tabouret-Keller (1985: 26) were the first to advance the hypothesis that the Portuguese were, at the outset of their colonial venture, already familiar with the idea of a pidgin thanks to the experience of *Lingua Franca* in the Mediterranean. There are different interpretations regarding the evolution of a Portuguese pidgin. Naro (1978) suggests that a *lingua de preto* (‘language of the blacks’) existed in West Africa as early as 1455 and served as a model for the development of a simplified Portuguese (or Portuguese pidgin). Goodman (1987) claims that pidgin Portuguese could have arisen in a number of places; there is evidence of pidgin being spoken in Cap Verde and São Tomé around the same time. It seems clear that some simplified Portuguese was spoken in West Africa between the Portuguese *lançados* and their African wives and that Afro-Portuguese did influence the development of a Portuguese pidgin used by European Portuguese later on (Holm 1989). Since the Portuguese in Asia were typically followed by large numbers of African slaves, servants, and soldiers, it is likely that features of such a pidgin would have arrived in the Asian-Portuguese settlements.

More recently an interesting variation of these claims has been put forward. According to Clements (1992), a type of foreigner talk (FT) had developed throughout the middle ages (fifteenth century) in areas first conquered by the Portuguese and then resettled by foreign people. When the Portuguese came into contact with new populations they could avail themselves of simplification strategies based on the experience of this FT. These include a reduced morphological system compared to Portuguese, simplified pronominal system, lack of articles, and zero copula. Among the origins of FT is also contact with African languages, which may have provided some of these features such as zero copula.<sup>11</sup> What is crucial is that these features are traced by Clements (1992: 86) to the language used by Indian priests in Korlai, one of the settlements of the *Estado da Índia* that was, like Chaul and Goa (and later Melaka), the source

from where Portuguese traders came to settle in Macau. This indicates some continuity in the evolution of Asian Portuguese, which also explains the strong retention of Malayo-Portuguese features in Makista, and the relatively limited influence of Sinitic.

### 3.4.1 *Asian Portuguese vernaculars*

In the early seventeenth century, a community of almost 8,000 Christians, including 300 Portuguese, had arisen in Melaka from intermarriages between Portuguese men and local (Malay and Indonesian) women. The Portuguese presence in Melaka lies at the origin of several contact languages referred to as 'Malayo-Portuguese Creoles'. These languages evolved from a mixture of Portuguese and Malay, not however in their standard or literary forms but rather in their colloquial varieties (Ansaldò and Matthews 2004). It is quite likely that Malayo-Portuguese varieties may have been influenced by Portuguese-based Creoles developed in West Africa or India (Hancock 1975; Clements 1992). For example, tense, mood, and aspect (TMA) markers are fairly similar across Asian Portuguese varieties: the future marker *bai*, the perfective *ja* (from Port. *já* 'already'), and a negative future marker *nadi* are found in Malayo-Portuguese (Holm 1989; Thurgood and Thurgood 1996), in Indian-Portuguese varieties of Cochin, Sri Lanka, and Korlai, and in Makista (Ferreira 1996: 242–3; Ansaldò and Matthews 2004). Some features may have an even wider distribution, such as *bai*, found in Melanesian pidgin. One well-known Malayo-Portuguese Creole which used to be spoken in Melaka is Papia Kristang (PK, lit. 'language of the Christians'). As noted above, after the Dutch takeover of Melaka, PK remained in use as a lingua franca between the Dutch and the locals. It is still spoken today by approximately 1,500 people (Holm 1989: 291). The term *Papia Kristang* is sometimes also used to refer to Makista (see chapter 7). As noted in Baxter (1996), the connection between Makista and Asian Portuguese is very clear. In the sound system, features such as weakening of Portuguese nasal vowels, diphthong/monophthong alternation in stressed syllables, and absence of final /r/ are all found in Makista as well as Melaka Creole Portuguese (Baxter 1996: 325). In Makista morphosyntax, we see variable gender, number, and subject–verb agreement, also common in Melaka Creole Portuguese, as well as reminiscent of Cantonese, which lacks agreement altogether (Baxter 1996: 326).<sup>12</sup> Besides the TMA markers already mentioned above, the Portuguese vernaculars of Macau and Melaka also share the preverbal markers *ta* (non-punctual), *lo* (irrealis), and *kaba* (completive), commonly found in Malay varieties (Baxter 1996). It is clear, judging from cultural as well as linguistic elements, that the Indian, Malay, and Chinese settlements founded by the Portuguese were in close contact and shared a number of common traits in all aspects of life (see also chapter 7).



Apart from the fact that Chinese authorities kept the Portuguese at considerable social distance, the internal dynamics of a city-port like Macau are quite typical of a Portuguese Asian enclave. In order to gain a deeper insight into the sociolinguistically relevant aspects, I review the history of Macau in the following section.

### 3.4.2 *The Portuguese in Macau*

The city of Macau developed as a consequence of Portuguese expansion in the South China Sea. It was partly through the contacts with Indonesian–Chinese communities that the Portuguese realized the economic potential of the Far East. Between 1513 and 1519, Portuguese traders visited the area of the Pearl River Delta and established bases for the purpose of trade. The Portuguese were unique among Western powers in being the first to gain access to the Japanese market, perhaps through the sale of firearms to the Japanese. Through their activities in China and Japan they were able to set up the most lucrative trade of the whole empire: the Japan trade. Macau thus became a central node for the trade of spices from the Southeast, silk from China, and silver from Japan, which soon eclipsed all other trade under Portuguese influence (Newitt 2005: 147). Such was the importance of the trade that Macau's growth in political power came to rival that of Goa, the centre of the Indian–Portuguese empire, the *Estado da Índia* (Boxer 1953; Newitt 2005).

Unlike the colonies of Goa and Melaka, the city-port of Macau developed without explicit authorization from either Portugal or China. Its existence was recognized after the fact by Portugal and was tolerated by the Chinese. By 1586, Macau was officially recognized as a city by the Viceroy of Portuguese India. Portugal itself did not maintain a significant presence in Macau's history, since, by the time of the establishment of Macau, the kingdom had lost its independence to the kingdom of Spain. True European-Portuguese people were indeed few in Macau; usually, when referring to the Portuguese, historical sources really mean people of Portuguese origin settled in Asia, predominantly from Goa and Melaka (see Amaro 1994; Batalha 1994; Teixeira 1994).

Because of the segregation of foreigners imposed by Chinese authorities (discussed in chapter 8), the social dynamics of a city like Macau differed from that of Melaka.<sup>13</sup> Within the fortified city of Macau lived the foreigners, including the Portuguese and individuals of Indian, Arab, African, and Philippine origins; the Chinese population lived outside the city walls, and had to leave the inner city by nightfall.<sup>14</sup> In the beginning, Chinese authorities regulated all contacts between foreigners and Chinese: *Porta do Cervo* (lit. 'gate of siege') erected by the Chinese in 1573 functioned as a border point. This gate was opened twice a month to allow fresh food supplies into the city, and constituted the lifeline of the population of Macau. The foreign population of Macau appears to have been relatively stable over time. In the sixteenth century, there



were approximately 500 to 600 'Portuguese' merchants. By 1635, António Bocarro, the general archivist of Goa, estimated that there were 1,700 *casados* in Macau, i.e. married Christians, of which half were 'white' and half were 'black'. Such a classification was common throughout Portuguese Asia and distinguished between the married Christians of Portuguese origin, mostly of mixed ethnic descent, and the married Christians of indigenous origin (Bethencourt 2005: 116–7). In that century there were also approximately 5,000 slaves in Macau. The number of Chinese living outside the city walls was steadily on the increase (there were none in the area before the arrival of the Portuguese): between ten and twenty thousand reportedly lived there in the seventeenth century. This increase in numbers testifies to the growth of Macau as a trading power and indicates a relaxation in the official controls imposed by Chinese authorities; later, with the expansion of the Canton trade (see chapter 8) and the defeat of the Chinese in the Opium Wars (1850s), the Chinese authorities lost control over relations between Macanese (and other Westerners) and the local Chinese population.

A Japanese community is also attested in the history of Macau. Its origin was in the port of Nagasaki, where a Luso-Japanese community had developed through intermarriages between Portuguese and Japanese women converted to Christianity. When the Japanese started persecuting Christians in the early seventeenth century, the community relocated to Macau. By the nineteenth century the 'Portuguese' contingent of Macau had increased to 900 people, with many additions from other Portuguese ports in Asia, and the Chinese population probably exceeded 50,000 (Amaro 1994; Teixeira 1994; Baxter 1996: 323).

Until the late nineteenth century, no Portuguese women accompanied the Portuguese traders, sailors, and administrators to the Asian settlements. At the same time, contact with and marriage to the Chinese was prohibited in the early decades of the colony by the Chinese authorities. Because of this tight separation between Chinese and Portuguese in Macau, the official wives of the Portuguese in Macau were seldom Chinese in the early phases. Instead, the female population of Macau came largely from other colonies in which Portuguese settlers could find wives, such as the Indian colonies and Melaka and Goa, and were predominantly of Asian-Portuguese, Malay, Indian, and Japanese origins (Boxer 1965; Amaro 1994; Batalha 1994; Teixeira 1994). It is, for example, recorded of several Portuguese settlers in Macau that they were married in Melaka to women of Malay and Indonesian origins, often of low social extractions. Women were also brought from the Portuguese community of Nagasaki in Japan.<sup>15</sup> Official records did not, however, take into account the presence of illegitimate offspring and local concubines (often referred to as *bichas*, a derogatory term literally meaning 'worms'); concubinage involving the Chinese was in fact common practice as attested in the chronicles of Peter Mundy (mid-seventeenth century), who lists the presence of many Chinese girls in the household as the one peculiar 'property' of Macanese heads of family (Amaro 1994: 25). Indeed, Cabral

(1994: 230) lists 'concubinage' as one of the cultural traits that distinguished Macau from the rest of the empire. We also know that the numbers of illegitimate children were high (Teixeira 1994: 100). It is a fact that Macau was from early on known as 'Paris of the Orient', a promiscuous society where men could openly keep several mistresses in or outside the household, a common practice in the higher strata of Chinese society.<sup>16</sup>

### 3.5 Summary

To sum up the sociolinguistic overview in this chapter, when we look at the evolution of vernacular varieties of the Monsoon region, it is clear that we must consider three language groups as critical 'players' in the region, in this order: spoken Malay and Indonesian varieties (Java Malay, Javanese, Sundanese); southern Sinitic (Hokkien, Cantonese, Teochew); and Asian Portuguese. (In addition, though not a focus of this book, Tamil also surfaces here and there as an early influential language of trade, most likely a leftover of the powerful Chola empire that had traded intensively with Sri Vijaya between the ninth and thirteenth centuries (Sastri 1976).)

A few points should be borne in mind regarding these critical players. When we talk about Malay and Indonesian influence, the diversity of features that we actually refer to is enormous, considering the typological diversity found in Indonesia alone, as well as the diversity of registers and special codes mentioned in this chapter. The contact was seldom, if ever, between literary or standard varieties, and we have to bear this in mind when engaging in detailed linguistic analysis. Further, as we have seen in this and the previous chapters, the most extensive, prolonged, socio-historically significant, and well-documented instance of contact is undoubtedly the one involving Hokkien (and other Chinese) traders based in the Southeast Asian region from the fourteenth century onwards but very likely earlier than that. The role of Sinitic in the evolution of a number of contact languages can therefore not be denied; but it should not be exaggerated either, because, as we will see in chapter 7, Chinese traders too relied on Bazaar Malay varieties in their everyday communication. In this sense, this is a history of contact languages in contact, and I will try to apply this to the development of Bazaar Malay and Makista in chapter 7.

In conclusion, chapters 2 and 3 have sketched the geographical, political, and sociolinguistic environment of Monsoon Asia, defined as peninsular Southeast Asia, the South China Sea, and the eastern part of the Indian Ocean. It is within this ecology that the processes of contact language formation described in the book take place, and it is therefore only in relation to the salient ecological factors that these processes of language evolution can be understood. This history has focussed on Southeast Asia, in particular Melaka and Java, as the cradle of those encounters that would shape the

history of the region. The main protagonists in this history are, on the one hand, Chinese traders and, on the other, Malay and Indonesian people, in particular women, slaves, servants, and concubines. The early and influential presence of merchants from Southern China, as well as the political ties between China and the Malay/Indonesian world, are solid historical facts. The city-ports were the socio-political environments in which the encounters were cemented. In these environments, Chinese and Malay/Indonesians were not alone, considering the significant presence of Indian and Arab communities, among others. Looking at the social structure of the city-ports increases our understanding of the fact that we are clearly looking not at simple bicultural and bilingual encounters, but at complex interactions of a multicultural and multilingual nature. As we can see from the political systems that emerged in the city-ports, the combination of exogenous elements resulted in the creation of new political and economic systems unique to the ecology of Monsoon Asia. The main thesis presented in this book is that the evolution of new communicative systems is just another aspect of this pluralist attitude towards a new, emerging environment and can be treated as one aspect of new cultural formation, which required new laws, new beliefs, and new languages. While it is historically plausible to place Sino-Javanese encounters at the origins of such development, we must acknowledge the role that other elements played in the formation of the new communities. The role of Islam contributed to the uniqueness of the city-ports and offered cultural traits that were neither Chinese nor Malay in origin, adding a significant element of hybridization to the new culture. Islam also offered a vehicle of transmission of many cultural traits to remote corners of the region, including linguistic traits. Conversely, the political transformations taking place in imperial China, emphasized by the arrival of Western powers, contributed to the relative marginalization and subsequent indigenization of many communities, strengthening their new identities and leading to linguistic stabilization (see chapters 6 and 7).

Though population admixture occurred to different degrees and in different patterns across environments, linguistic admixture clearly occurred in the evolution of languages in the ecologies described in chapters 2 and 3. It is in these ecologies that human cultures converged and diverged; the types of social dynamics underlying them, as well as the behavioural patterns that emerged from them, must be seen as utterly normal in the sense of their being frequent, and therefore unmarked. It follows then that the linguistic patterns of such ecologies must also be utterly normal and unmarked, and as such should constitute a robust feature of any realistic theory of language. Following this account, I propose that language contact, and contact language formation, must determine our theoretical generalizations on language; we should not struggle to find a proper place for them. Bearing this in mind, the crucial question is: what does a theory of language evolution look like in which language contact

plays a central role? The answer to this question has philosophical, theoretical, and methodological implications discussed in chapters 4 and 5.

## Notes

- 1 The role of Sinitic is also discussed in detail in chapters 7 and 8.
- 2 Literary Malay was later to be adopted by the Dutch and the British as the language of local administration, and, in the beginning of the twentieth century, standardization of this variety was started as a means of modernizing the language.
- 3 The representation of Malay data follows conventional Malay scholarship.
- 4 The inherent diversity of Bazaar Malay varieties and the time-depth involved in their evolution make it very difficult to arrive at plausible reconstructions of their grammars. The set of core features presented in Adelaar and Prentice (1996) as 'Pidgin-derived Malay' features is probably the closest we can get in terms of a structural profile.
- 5 This is first introduced in Adelaar and Prentice (1996: 675) as 'Pidgin Malay Derived' varieties.
- 6 These are listed in approximate order of dominance.
- 7 Collins (1984: 155), for example, notes that terms such as *Creole* and *pidgin* retain a colonial viewpoint.
- 8 These are all dangerous implications that, I believe, should not be encouraged or propagated in linguistic research, but rather critically addressed and deconstructed.
- 9 In my reading of Adelaar and Prentice (1996: 675), there is no clear reference to the notion of a Malay pidgin having existed at any point in time; instead, their notion of PDM varieties implies the existence of lingua francas that shared a set of common grammatical features.
- 10 The Dutch language in contrast left a very weak trace in the region, which can be explained by their policies of segregation and the reliance on Malay and Portuguese in Dutch colonies such as Sri Lanka (see chapter 6).
- 11 Alleyne (1971: 184) had 'predicted' a historical connection between the West African Portuguese settlements and their Asian territories.
- 12 According to Clements (1992), this could be an early feature left over from the pidgin Portuguese which played a significant role in the development of Portuguese Creoles.
- 13 Portuguese enclaves in Asia in general can be characterized by a relative social separation from other local communities, in particular due to the clear religious association with Christianity.
- 14 The only exception was a small group of Hokkien fishermen, allowed to live within the city walls. They helped the Portuguese through fishing, carpentry, and other skilled labour, and may have had an early influence on the Sinitic component in Makista (Ansaldò and Matthews 2004).
- 15 Little obvious influence from Japanese can be seen in Makista, although Batalha (1977) notes a few loanwords such as *auabi* from Japanese *awabi* 'abalone'.
- 16 Chinese women were present as housemaids and concubines from the seventeenth century onwards, when heavy migrations from Southern China to Macau (and Hong Kong) took place (Ansaldò and Matthews 2004). As mothers and nannies involved in bringing up children, such Chinese women contributed to creating a multilingual ecology from an early stage.

## 4 Methodological issues in the study of contact languages

---

*The consistency condition which demands that new hypotheses agree with accepted theories is unreasonable because it preserves the older theory, and not the better theory. Hypotheses contradicting well-confirmed theories give us evidence that cannot be obtained in any other way. Proliferation of theories is beneficial for science, while uniformity impairs its critical power. Uniformity also endangers the free development of the individual.*

Paul Feyerabend (1975: 36)

What is a theory of something? What does a theory aim to do? The word *theory* comes from Greek and originally meant something like a way to look at things, a point of view. A theory is basically an idea, usually centred within one specific worldview, or a philosophical idea, about the nature of a given aspect of the world. Depending on the basic ideas that we bring into the experience of any event, we influence and, to some extent, limit, or narrow down, our field of enquiry. Humans are naturally curious and like to hypothesize about the world. Most academics believe that creating hypotheses and testing them can further our understanding of the world we inhabit. This is in itself already only relatively true; it holds within one specific worldview, the ‘scientific method’, influenced by Western philosophy since the time of Socrates. In Buddhist philosophy, for example, rational enquiry reveals nothing about the true nature of existence because it can only touch on superficial, illusionary aspects of reality. Real knowledge in Buddhism can only come through introspective meditation (and in Greek, ‘theory’ could indeed also refer to ‘contemplation’). So we should really see theories as culture-specific and time-bound ideas that humans develop in attempts to increase their knowledge of the world. According to the scientific method, theories should be tested and discarded, or modified, once they reveal themselves useless or erroneous. Crucially, the testing can occur either when attempting to explain new things, or in the reinterpretation of old things, as I will explain below.

In linguistics, as in other fields of enquiry, there are many, at times contradictory, ideas of language, its origins, its function, and its ‘mechanisms’. At times one may get the sense that, in linguistics, there are more theories out there than actual evidence to support them. Until recently, for example, we had

many more abstract ideas about ‘universal’ grammar, i.e. the alleged initial structure of all human languages, than actual knowledge of the grammatical structure of many of the languages of the world. Even today, we only have grammatical information about 20 per cent of the world’s 5–6,000 languages, and of that only a small subset is actually reliable. However, theories about the alleged universality of linguistic structure abound and proliferate. In what follows, I address some of the more relevant linguistic ideas for the purposes of this study, in particular ‘theories’ about how new languages are formed, how language change happens in history, and how contact between languages affects their grammars. For now, let us intuitively explore some ideas that may help us to understand communication between speakers of different languages.

Imagine that you find yourself in a community whose language you do not speak. You find that the local inhabitants do not speak your language, and you need to take care of many things, such as finding food and shelter, arranging transportation, and asking for directions. But every time you try to communicate, you fail to be understood. What do you do? Anyone’s impressionistic observations, based on experiences of being a foreign element in different communities, suggest the following to be the most common patterns: (a) you try to impose your own language; (b) you engage in a communicative trade-off: you learn some, you teach some, you make some up until communication is functional; (c) you try to learn the new language.

Strategy (a) is usually unproductive, but common for members of monolingual speech communities with long, normative traditions. Strategies (b) and (c) are examples of what you can observe especially in seasoned travellers and explorers, and on a larger scale in many non-Western cultures or traditional environments where people from different linguistic backgrounds come together on a regular basis, i.e. markets, ports, border towns, etc. For example, in many markets of Southeast Asia, you will hear people using words and phrases derived from Malay, Chinese, and Indian languages as well as English, all in the same conversation; this is the result of a long communicative trade-off that has been happening for a long time, already described in [chapters 2 and 3](#). This is not unlike being a tourist in a foreign country, where you find yourself wanting to bargain over the price of that necklace: you may try to use the words of the local language to speak the numbers – surely your pronunciation will be influenced by your own language – you will point to things and say the name for them in your language, and ask ‘how to say’ that in the local language. The merchant will be keen to remember how you refer to those same items just in case another tourist with the same language as yours comes around. If you have studied any other languages, and if you think they might be helpful, you will tap into that knowledge to see whether that improves mutual intelligibility. The longer you find yourself interacting in that community, the more features will be exchanged, especially if you are together with other speakers

of your language who also contribute to the exchange. This is a very simple yet extremely common example of how language contact occurs. Once you have some more fluency, you will find yourself using all the phrases and words you can, and when you run out of options you will fill in the blanks with your own language. This is called code-mixing, and is one of the most natural communicative strategies in multilingual environments.

My point is that in naturalistic linguistic environments, people use languages spontaneously to achieve communication. If these languages are new, they will trade features, adopt and adapt them until mutual understanding is possible. We should not conceive of speakers in multilingual settings as prisoners of grammatical rules, rigidly frozen around their own language; speakers, especially when equipped with more than one language, are not constrained by innate linguistic patterns that prevent them from adopting other ways of speaking (Yip and Matthews 2007) – apart from perhaps the sounds they are unfamiliar with.

The theoretical idealization I adopt in my approach is that speakers are creative users of the linguistic resources around them, and in contact language formation we find the evidence in support for this. Contact language formation is therefore, in different degrees, the natural, ‘normal’ outcome of multilingual encounters in various socio-historical contexts (Mufwene 2001), in line with a view of language acquisition as a reconstruction process (Hagège 1993). However, for the longest time, contact language formation has been seen through the lens of idealizations in which linguistic purism and linguistically innate features combine, leading to the implicit assumption that linguistic structures hold speakers captive and prevent them, or inhibit them, from engaging in creative communication, unless in extraordinary situations. From this, a whole array of special or exceptional ideas developed into theories to account for the formation of new language as a rare, extraordinary phenomenon.

#### **4.1 The ideology of theory**

In social theory, the relation between language practice and the political-economic sphere has received substantial attention, especially since the work of post-structuralists such as Foucault (1972, 1980) and Bourdieu (1977). In these works, a particular aspect of theorizing is highlighted in which linguistic (and other) models come to influence researchers to the point of dominating the academic discourse of certain institutions (a ‘paradigm’ in the Kuhnian sense of a set of practices by which a scientific community is defined, Kuhn 1970). Silverstein (1979: 193) aptly defines ‘linguistic ideology’ as “sets of beliefs about language articulated by users as a rationalization or justification of perceived language structure and use”. In such a definition, the link between subjective or culture-specific ideas and observation of reality is made explicit.



As noted in Kroskrity (2000: 6), such a link may be an accepted truism of cultural anthropology (see the work of Geertz, especially 1973, 1983) but is not necessarily acknowledged in all linguistic fields. The field of linguistics, which is still dominated by the positivist philosophy, seems however to ignore post-structuralist thought (Kroskrity 2000). Boas (1911) and Bloomfield (1944), for example, raise the knowledge of the linguist to a level of expertise that transcends the link between subjective perceptions and objective knowledge and becomes ‘scientific’; as such it is superior to, for example, native theories of language. Related to this is the impact on linguistics of the structuralist turn pioneered by Saussure and Levi-Strauss (Caws 2000). The introduction of a distinction between the psychological (*langue*) and the social (*parole*), fully exploited in the works of Chomsky, leads to an abstract notion of grammar as detached from usage, context, and history. This is still very much the case in studies of language structures where the body of current knowledge is taken to constitute an objective set of facts through which new enquiry can be filtered and classified, and where system-external factors, such as social, political, and ideological aspects, are considered peripheral to a scientific theory of language. Below, I look at the fundamental ideas of the field of historical linguistics and contact linguistics. Following models such as those of DeGraff (2003), I illustrate the connections between history and ideology in theory formation and show how failing to understand the origins of our theories leads us to the construction of questionable explanatory frameworks.

#### 4.1.1 *Historical linguistics and its foundations*

The foundations of the field of historical linguistics today can be traced back at least to the field of Indo-European studies that developed in particular during eighteenth- and early nineteenth-century Europe. This field was steeped in Indo-Germanic philological traditions and developed through the comparative study predominantly of Sanskrit, Greek, and Latin. One of its aims was to establish genetic relationships between languages of Eurasia (Bopp 1857; Schleicher 1861); this exercise culminated in the reconstruction of Proto-Indo-European, the *Ursprache* (‘proto-language’) of European people, as well as in a series of conjectures about the land of origin of all European populations (see Renfrew 1987). In this historical phase, a powerful paradigm can be found in popular interpretations of Darwin’s evolutionary theory. The myth of the Indo-Europeans, when interpreted within an evolutionary framework in which history is seen as some kind of social progress, was open to easy manipulation by social currents of nineteenth- and twentieth-century Europe, and fomented movements such as Nazism (Arvidsson 2006). In linguistics, this can clearly be seen in the development of the *Stammbaum* (or genetic tree) model of language speciation (Schleicher 1863).<sup>1</sup> Note that in the intellectual environment



permeated by romanticism and nationalism of eighteenth- and nineteenth-century Europe, a strong equation between a people and a language was required (Coulmas 1988; Fishman 1989; Blommaert and Verscheuren 1992; Errington 1992). Establishing the historical pedigree of a language was primarily a political and ideological exercise that carried racial implications (Eckert 1980; Aarsleff 1982; Harris 2004). In fact, implicit in the *Stammbaum* theory is the structuralist notion of languages as discrete entities that can be passed from speaker to speaker, or from generation to generation, as entire systems. The nature of such transmission is crucial to the idea of linguistic purity: if the linguistic code is transmitted as a whole from one population to another, without external interference of other linguistic codes (i.e. language contact), it is considered a 'normal' process which confirms a genetic link between parent and daughter language. In accordance with this, the types of structural changes that may be observed between generations will also be considered normal types of language change. This model has left us with a legacy of concepts that, to this day, tend to be regarded as somewhat universal and are thus invoked to account for the evolution of all 'normal' languages, independent of their historical and socio-cultural background. Purity is compromised by evidence of more than one language in the transmission process: according to this, mixed languages are impure and genetically isolated.

One problem lies in the fact that the *Stammbaum* is at times misinterpreted as offering an *account* of language diversification rather than a *representation of the outcome* of a speciation process (Mufwene 2007: 64). Another problem lies in the fact that the model has shortcomings, for example, in that the influence of contact in language speciation is disregarded or treated as secondary, and the role of horizontal, as opposed to vertical, diffusion of traits is minimized. Schmidt (1872) proposes the *Wellentheorie* to account for horizontal (areal) propagation of features, a theory largely ignored in historical linguistics until the advent of areal typology, where, for example, Haspelmath (1998) and Heine and Kuteva (2005) argue that Europe can be seen as a linguistic area, in which contact and horizontal diffusion, rather than genetic inheritance, can account for common features. There is also substantial work on Australian aboriginal and other languages of the world suggesting that their evolution cannot be captured in a vertical model, because diffusion and convergence played a greater role than inheritance in the evolution of these language families (e.g. Durie and Ross 1996; Dixon 1997; Aikhenvald and Dixon 2001; Mufwene 2002a). Diffusion can thus be seen as a type of horizontal transmission in the evolution of languages, parallel to the notion of horizontal acquisition invoked as the dominant model of transmission in Mufwene (2001).

Because of the cultural dominance of nationalist ideas developed in the European tradition of the time, the observations derived from the investigation of Indo-European languages in this environment were eventually promoted to

being accepted as general patterns of change, which languages undergo in ‘normal’ circumstances. Here the work of the Neogrammarians, in particular, laid the foundations for a number of ‘natural’ laws of change: as already suggested in Dahl (2004), the notion of ‘natural’ in the Neogrammarian model is likely to have resulted from a biased analysis of the situation. Take the correlation between the reductive effects of phonological change on unstressed syllables and loss of case marking. According to Blake (1994), it is likely that the disappearance of nasals in Latin would have reduced word-final *-m*, causing the loss of accusative marking. But, as shown clearly in Dahl (2004: 169), maturation processes in languages abound with examples where structural complexity is maintained even though the weight of the phonetic expression decreases. In other words, another possible outcome for the Latin example would have been morphologization of the feature of nasality. Thus the notion of ‘natural’ in this context should really be interpreted as ‘familiar’, as it is based on lessons derived from the observation and description of Indo-European languages.

We should therefore realize three important things: (1) theories result from the ideology of the time that produces them; (2) a model developed in one specific empirical domain may not be able to capture facts about different empirical domains; and (3) the lessons derived from a model are easily misconstrued as universal or ‘normal’. In this sense we can say that the field of (historical) linguistics suffers from a certain Eurocentric bias in which monolingualism and normative transmission, which allegedly ensure ‘purity’, underlie all theoretical assumptions of ‘normality’, what Hutton (1998) refers to as ‘mother-tongue fascism’. In this intellectual climate, monolingualism always marginalizes the complexities of multilingual societies, which have been peripheral to the construction of linguistic theories (Kachru 1996; section 4.3 below and chapter 5 of this volume). Considering that, to this date, Western ideas and ideology still dominate the economic-political and therefore academic discourse, as can be seen in the prestige of Western-style institutions and Western-trained academics in most fields, it is safe to say that not much has changed in the past century. From the Indo-European paradigm of historical linguistics, a putative correlation between types of change and types of structural outcome emerges as follows. If a language and a people have a normal history, this is reflected in the linguistic history showing continuity, and therefore some degree of purity, across generations. If, at some point in time in the history of a population, important changes are noticed in the linguistic codes of different generations, some catastrophic socio-historical event must be assumed, and purity is compromised.

#### 4.1.2 *Creole exceptionalism*

The correlation between language admixture, restructuring, and socio-cultural upheaval leads to the development of a particular theoretical paradigm in Creole studies, identified by DeGraff as ‘Creole exceptionalism’ (2001; see also

Mufwene 2001; Ansaldo and Matthews 2007; chapter 6 of this volume). Creole exceptionalism refers in particular to the position that Creoles are structurally impoverished variants, or degenerate offshoots, of their European norms. This, according to DeGraff (2001, 2002, 2003, 2005), is an explicit case of *minoration linguistique* resulting from the historical and ideological assumptions of the seventeenth to nineteenth centuries already presented above. In Creole-related writings of the colonial era, ‘programs of perception’ were congruent with race theories that provided philosophical justification for the New World slavery. Creole theories, DeGraff explains, were ordered by Europe’s ‘normative gaze’ vis-à-vis the non-white and non-Christian world. In 1894, Poyen-Bellisle describes, in a University of Chicago dissertation, how slaves try to imitate ‘civilized man’s speech’ but fail because they allegedly do not belong to the same race, and their vocal tracts and their lips are different (Poyen-Bellisle 1894). The ideological bias associated with the idea of purism can be detected in the age of discovery in the race- and evolution-related myths that result in the congruence between the evolution of races and that of languages typical of nineteenth- and twentieth-century historical linguistics. Thus Meillet (1948) describes Spanish Creole or French Creole as varieties of Spanish and French *deprived* of almost all their grammar, *weakened* in their pronunciation, and *reduced* to a small lexicon (my emphasis). The power of the culture = race = language equation is still alive today, as Creole genesis is considered to fall outside the scope of the Comparative Method, given that Creoles (and other types of mixed languages) are regarded as non-genetic ‘orphans’ outside the family tree of human languages (Mufwene 2003).

## 4.2 Multilingualism and transmission

Although purity plays an important role in historical linguistics, the notion of language contact as a fundamental process of any environment of language transmission has at the same time found recent support in ecological views of the evolution of languages, and notions such as ‘pure’ and ‘normal’ have become increasingly relativized in approaches to language contact and change (see in particular Croft 2000; Mufwene 2001, 2008; Ansaldo 2004a; Ansaldo, Matthews and Lim 2007). In these approaches, the following truisms that characterize many studies of language change and language contact have been criticized and deconstructed:

- There is a distinction between ‘pure’ and ‘hybrid’ languages.
- Normal transmission guarantees genetic relationship.
- Radical contact varieties have a non-genetic origin.

In recent work on the ethnography of speaking, the universal validity of linguistic norms across speech communities has indeed been questioned (Ochs and Schieffelin 2001; see also Bloomfield 1933). It has been argued

that the traditional sociolinguistic position of structured variation and diversity must give way to a more complex and flexible notion of speech community (Rampton 2001), which falls in line with late modernist views that suggest that the reality we aim to capture is “much more fluid, heterogeneous and under-patterned than anything sociologists have tried to grasp intellectually in the past” (Bauman 1992: 65). In this sense, the view that an entire single set of rules can be transmitted unaltered from generation to generation stems from a view of society as linguistically homogeneous.<sup>2</sup> This is the monolingual view of language and socialization already criticized, among others, in Skuttnab-Kangas (2000) and Leather and van Dam (2003). Non-Western societies tend in many cases to be linguistically heterogeneous as we do normally find at least dialectal variation in any given speech community. As I discuss in [chapter 5](#), this initial variation is quite likely to be at the origin of the ‘internally motivated’ types of language change, i.e. where multilingualism and language contact is less prominent (Whinnom 1971; Croft 2000; Mufwene 2001). This approach calls into question the notion of a dichotomy between ‘normal’ and other types of transmission.

Before qualifying what we mean by normal transmission, we need to look at communicative environments in a global perspective. Multilingualism and bilingualism are widespread and still account for the majority of the world’s population (Edwards 1994; Auer and Li Wei 2007). The majority of Africa’s population is bilingual or multilingual (Wolff 2000), and, in all of sub-Saharan Africa, several languages are used in education at the same time. Asia is largely multilingual in practice as well as in education, in particular mainland and peninsular Southeast Asia (Goddard 2005). Even China, often conceived of as a monolingual Mandarin-speaking country, is in fact at least bilingual, as Mandarin is a standard variety usually acquired later in life, and other mutually unintelligible languages are spoken by a majority of the population as native vernaculars (Norman 1988; Ansaldo 1999). India’s multilingualism is so pronounced that it has been regarded as a problem to solve by educationalists and politicians (see e.g. Annamalai 2001). Papua New Guinea is the multilingual country per definition, where linguistic diversity is the highest in the world (see Foley 1986; Smith 1995). Even in regions where the impact of Western culture has prompted language death and shift to monolingual systems, we still find multilingual areas of great diversity, such as the Amazon region, or else widespread bilingualism, as in other parts of South America. It is thus easy to see that monolingualism is a condition mostly affecting Western European countries and a few other similar systems, but is not the natural status of a majority of societies around the world. If we factor in the rising rate of language loss, we must conclude that in the past the multilingual nature of human society must have been even more pronounced. Add to this the fact that schooling and standardized languages are very recent inventions that have yet to spread to

many developing countries. This implies that, in most regions of the world, language transmission is most often not tutored nor strongly enforced by societal requirements of accommodation to the ideal norm.

In multilingual speech communities, a significant degree of linguistic variation and interference between languages must be expected, as we can see in the case of Singapore. Even if the four officially recognized languages – English, Mandarin, Malay, and Tamil – are clearly compartmentalized through the education system, supported in school and ideologically enforced by the government, mixed languages such as Bazaar Malay and Singlish are thriving (see [chapter 7](#)). Imagine how much more frequent language admixture is in multilingual societies where institutional support of the different languages is not found. As noted in Le Page and Tabouret-Keller (1985), ‘imperfect’ acquisition of language is particularly natural in non-scholastic environments, and is still very frequent in today’s societies, in which developing communicative skills is more a product of communicative efforts than a planned, regulated activity. In such communities, when multilingualism is high, it is difficult to imagine that a ‘pure’ set of linguistic features – even if it existed – derived from one single system could be passed on in its entirety without interference from other languages. This opens up the possibility of significant adaptation and admixture in the transmission process. Hagège (1993), for example, reminds us that language transmission always involves a combination of inheritance and creation; therefore innovative replication of features is a fundamental factor behind language change (see also Meillet 1948; Lass 1997; Croft 2000; Mufwene 2001). In such contexts where heteroglossia and informal transmission combine, variation and code-mixing are normal linguistic practices of a community, ensuring a more creative use of language than the one witnessed in monolingual environments. The fact that multilingualism and code-mixing are often problematized and regarded as potential causes for cognitive impairment and/or social stigma is the result of the advent of European standard languages (in the seventeenth to nineteenth centuries); but we should not forget that it took several hundred years to turn most of Europe into prevalently monolingual societies. Until at least the sixteenth century, code-mixing between vernacular languages and Latin, for example, was the norm in most parts of Europe (see Auer and Li Wei 2007: 2–3). Nonetheless, we will see that monolingualism has been taken for granted in much theorizing about language, and that multilingualism has been consequently associated with exceptional developmental scenarios; in the light of the discussion above, this must be seen as a biased assumption.

#### 4.2.1 *Contexts of transmission*

Once we realize that monolingualism is not a normal condition of a community of speakers, we obviously need to call into question the related idea of

a ‘normal’ type of transmission. The dichotomy between normal and abrupt transmission is illustrated in the following quote:

We have argued that the most extreme products of linguistic interference [read: language contact] are not genetically related to any of the languages that contributed to their lexical and structural systems, because they did not arise from a process of normal transmission – that is, the transmission of an entire single set of interrelated lexical and structural features. (Thomason and Kaufman 1988: 200)

As argued in the work of DeGraff (section 4.1.2 above), the assumption of non-normal or exceptional events in the evolution of languages often leads to the postulation of extraordinary explanations that fall somewhat outside general linguistic theory. This can easily be seen, for example, in the fact that contact-induced change is still associated with less than full linguistic competence and therefore is seen as decay (cf. Rickford 1985; Jourdan 1991; Schieffelin and Doucet 1994). Languages that evolve in contact ecologies are therefore usually related to exceptional developments, whether structural, historical, or cognitive, as can also be seen in theories of Creole genesis (see also Ansaldo and Matthews 2007).

Though there is more agreement these days among linguists over the fact that contact-induced change is an essential and natural part of language (Mufwene 2001), some linguists are still struggling to come to terms with just how much change a language can undergo before it becomes a different language. As I show in chapter 6, regarding Sri Lanka Malay, it is unclear what such investigations can reveal, besides producing linguistic classifications that can be used – and are *misused* – for the construction of ethnic claims and racial discrimination (Ansaldo 2005, 2008).

Thomason and Kaufman (1988) discuss the various dimensions of language shift, interference, and contact in the evolution of languages. As illustrated in the quote above, even in this approach, we find a dimension of ideological bias in their definition of ‘not normal’ transmission. As long as just two languages are in contact, their argument goes, it is mostly possible to assess the degree of typological interference in the restructuring process. When three or more languages are involved, as is the case of most so-called ‘Creoles’ and many contact languages, it often becomes impossible to see the relationship between the input languages and the output grammar (unless strong typological congruence takes place between some of the varieties involved, which is also related to demographic factors). In these ecologies, the authors claim, competence in the native language would have been insufficient for proper communication (= exceptional explanation no. 1). A contact language would have been necessary and this would have led the children (if any were present) to grow up speaking the new emerging language as their vernacular, instead of their parents’ former vernaculars (= exceptional explanation no. 2). This, the authors feel, must be

a rapid process (= exceptional explanation no. 3). According to Thomason and Kaufman, the output, in these cases, is clearly not genetically related to the input (1988: 152). The evidence for this, they tell us, is found in the fact that ‘radical’ Creoles (e.g. Saramaccan) show very few European-derived features. What I want to point out here is the stacking of assumptions implied in this view that leads to exceptional explanations, which runs as follows:

- a. In multilingual contact environments, the typological relationship between input languages and output grammar can be obfuscated.
- b. Children in some of these environments may not learn their parents’ language(s).
- c. Therefore a new language may develop rapidly.
- d. Some of these languages show so few European features that they cannot be genetically related to any other language.

Therefore, Thomason and Kaufman conclude:

This removes all these creoles from considerations for genetic classification: they are not changed later forms of any parent language. They did not, in our view, arise through any sort of direct transmission from one speaker group to another, either by generation-to-generation/peer-to-younger-peer enculturation or by shift involving acquisition (perhaps with modification) of the grammar of the TL [Target Language]. Instead, they were created in various multilingual communities by the first generation of slaves. (1988: 152)

Let us focus first of all on claim (a): it is indeed true that contact between different languages can lead to a restructuring in which a typological shift takes place. This is, for example, the case of Sri Lanka Malay (SLM, see [chapter 6](#)), in which we see a language that preserves substantial Austronesian lexical items evolving into a grammar of Lankan type (Ansaldo 2005, 2008, 2009a). In these cases it is futile to discuss what the typological classification should be: the bulk of the lexicon is of one type, the bulk of the grammar of another. Genetic affiliation, however, can mean more than typological similarity. As I argue in [chapter 6](#), it is not necessarily the place of the linguist to assign genetic affiliation in these situations. In the case of SLM, the speakers perceive themselves as direct descendants of the Malays and Indonesians who came to Sri Lanka during the colonial era; their language is therefore also related to Malay (Austronesian). It would be difficult to claim that SLM is not a ‘changed’ language of its ancestor, most likely a Malay lingua franca, albeit changed to a high degree. But direct transmission is most definitely involved in the evolution of SLM, (e.g. Malay vocabulary), and shift through acquisition is certainly involved. How can any language arise otherwise? Therefore SLM is most definitely a genetic development, but of multiple ancestry (in the sense of Mufwene 2001). The role of the linguist, I suggest, should be to develop theories that are realistic enough to account for different types of speciation,



not to referee in a genetic contest which is, ultimately, an ideological, not an empirical, matter.

Claims (b) and (c) are related by the idea of ‘abrupt creolization’ as a means to counter the lack of a communicative system (see also [chapter 5](#)), which rests on the idea of a non-verbal stage in the history of a given community. These ideas, developed in particular in the work of Bickerton on Hawai‘i Creole English to account for the pidgin-to-creole development (Bickerton 1981, 1984), have been called into question. The notion of abruptness has been criticized as part of the accepted doctrine of creolistics for lacking empirical and theoretical support (Mufwene 2003: 274). Abruptness is often a consequence of a gap in our knowledge; the best example is Hawai‘i Creole English, which Bickerton had described as a classic case of abrupt formation, but has, through thorough historical documentation, been shown by Roberts (1998, 1999, 2000) and Siegel (2000, 2007) to have in fact emerged gradually over a couple of generations and from a pidgin that already contained a number of features found in the Creole. Alleyne (1979: 75) indeed stresses the bilingual or multilingual nature of early plantation societies, which would imply normal communicative interaction between different ethnic groups. Another case against abruptness is found in SLM: the fact that we can see the links between SLM and its input languages is due to a conflation of factors, including appropriate linguistic documentation of input varieties (at least two of them) and the fact that it is still spoken in a similar environment to the one of its formation. In other words, we are in a position to reconstruct the ecology in which SLM has evolved, which allows us to appreciate the following aspects of its evolution: (i) time-span: roughly three hundred years; (ii) multilingualism: at least three different languages involved and significant individual multilingualism; and (iii) typological pressure of the adstrate languages in the selection and replication process (see [chapter 6](#)). Without an almost complete reconstruction of its ecology, SLM may well have appeared to us as a radical Creole involving the typical exceptional explanations of Creole studies, such as (i) abrupt creation, (ii) non-verbal stage of formation, and (iii) independent developments ([section 4.2.2](#)). As Darwin (1859) pointed out, gaps and missing links in our evolutionary accounts are mostly indications of the limits of our knowledge.<sup>3</sup> Turning to claim (d), we should note that the degree of European retentions present in a contact language in which European varieties were involved is by no means a sign of a non-communicative stage of a community during which children have to tap into non-verbal knowledge to develop new grammar (which is what (b) implies). The amount of European features found in a contact language is likely to be the result of the degree of involvement that European varieties had in that particular ecology, the presence of non-European languages for inter-group communication, the willingness to learn the language of the ‘masters’, and other ‘external’ factors. As we will see



in [chapter 5](#), the problematic assumptions (a) to (d) can be overcome in a uniformitarian account of language evolution (Mufwene 2001), in which the processes of selection and replication of features within a specific linguistic ecology account for language change at large.

It is also important to acknowledge the role of acculturation in language genesis, as already advanced in Alleyne (1971) and Schumann (1978a), as well as Le Page and Tabouret-Keller (1985). As argued by Alleyne (1979: 96), since we all agree on the significance of social factors in language genesis, if we want to understand the process that leads to the development of new languages, we should first and foremost concern ourselves with the “reconstruction of the sociolinguistic situation at the time of [its] formation, with the structure of the communicative network, and the communicative needs of different sectors of the populations involved”. The validity of this statement has been confirmed in more recent studies in which contexts of formation are given serious consideration. These studies reveal a complex picture of environments where communication takes place that transcends the classic plantation scenario to include many more informal environments and social domains of interaction (Chaudenson 2001; Faraclas, Corum *et al.* 2007). The continuity of communication that we derive from the availability of appropriate historical documentation and understanding of the complex web of interrelated networks in the formation of Creole varieties renders claim (b) basically false. A non-verbal stage in which communication between generations is interrupted could only realistically exist in the case of a generation of orphan children on a deserted island.

The importance of the communicative network also shows the non-claim involved in (c). Thomason and Kaufman (1988: 153) suggest that the principles at play in contact-induced change are the same in bilingual and multilingual settings. Essentially, they argue that, when in a new linguistic environment, speakers will “make guesses about what their interlocutors will understand as they try to talk to one another. Those guesses that promote intelligibility will be the ‘right’ guesses. To begin with, the grammar of the emerging Creole will be a direct reflection of the shared ‘right’ guesses made by the shifting speakers”. If this is really true, how, we should ask, can this process of guessing happen faster in a multilingual environment (even more so if we were to assume lack of availability of ‘target’ languages) than in a bilingual one? Where many more languages and fewer obvious target systems are available to the speaker, it would be logical to expect a prolonged process of agreeing on the ‘right’ form, a slow process of societal focussing rather than a ‘rapid’ notion of Creole formation. As we will see later ([chapters 5 to 7](#)), an evolutionary perspective on propagation suggests that ‘speed’ is simply a reflection of the type and size of community in which language change occurs, as well as of the specific

conditions of language appropriation (such as being in a condition where no other lingua franca can be used).

#### 4.2.2 *Limits of classification*

An excellent illustration of the ideological nature of classificatory debates is the study of Old English and the degree of influence it received from Norman and Norse contact, presented in Thomason and Kaufman (1988: 263–331). As we can see in this work, a comparison of a list of features across different varieties and different eras of the history of English ultimately results in a matter of agreeing on external measurements. If we agree that a language can contain up to 25 per cent of material from other linguistic codes without having to be reclassified as a different genetic outcome, then even the most divergent dialect of English spoken in Great Britain can be seen as English, and therefore English is not a mixed language. It is obvious that the features we choose to investigate (versus those we leave out) influence the outcome. As pointed out in Mufwene (2003), establishing *a priori* the level of mixedness that a language cannot cross in order to be classified as genetically related to one of its plausible ancestors is inherently problematic. An example is the case of Singlish (or Singapore Colloquial English), typically considered a new variety of English or an ‘Asian English’ (Pakir 1991; Schneider 2003). This is based primarily on the recognition of the abundant lexical influence from English as well as the existence of grammatical material such as *wh*-questions and copula. However, as shown in Ansaldi (2004b, 2009b), *wh*-movement and copula are not so common in Singlish, which, more often than not, has in-situ questions and zero-copula. These features, taken together with topic–comment patterns, reduplication, and triplication for grammatical purpose, property verbs, aspect markers, ‘get’-passives, noun incorporation, and sentence-final particles, clearly show that Singlish is a language of the Sinitic type. But what is really the significance of such a classification? In fact, I believe that treating Singlish as an exclusively English, or an exclusively Chinese variety, would be equally misleading in terms of language evolution. The point I want to make is that categories are inherently static constructs and should not influence our understanding of diachronic processes. Hjelmslev (1938) remarked that all languages involve some degree of admixture, i.e. selection of features from different linguistic codes and recombination in a new grammar; and Mufwene (2003) notes that even some of the best documented cases of genetic classification followed an evolution in which language contact was abundant, as shown in the case of Romance languages (see Trask 1996).

If we agreed with the statement by Thomason and Kaufman quoted in the opening of the previous section, we might be tempted to think that the lessons derived from the study of contact linguistics might not be too significant

for the study of ‘normal’ language change. This has indeed been the case, as shown by the omission of Creole languages from typological studies of contact languages (see Aikhenvald and Dixon 2006: 7; and [chapter 1](#) section 5 of this volume). Following Mufwene (2003: 274), I believe that this is the result of the ‘dogma of intercourse-free speciation’, i.e. the assumption that ‘normal’ descent involves only one parent language. This has the unfortunate consequence of treating obviously mixed languages as anomalies that fall outside the central concerns of linguistic inquiry.

The field of traditional historical linguistics tends to focus on the mechanisms underlying internal change – i.e. changes that are believed to occur within the system by themselves without any human, social, or historical agency required in the description (Lass 1997; Lightfoot 1999) – based on the assumption that these can be isolated from a messy context, and therefore ‘properly’ analysed. On the other hand, the field of contact linguistics is more often than not concerned with external change, as it studies social and structural aspects of contexts in which two or more languages interact.<sup>4</sup> The difference between the two fields may lie in the fact that, in historical linguistics, the interest lies primarily in mechanisms of *inheritance*, while in language contact the interest is primarily in mechanisms of borrowing, transfer, and independent innovation. But, as argued, for example, in Heine and Kuteva (2005), there are many parallels between contact-induced change and other types of change, and the role of contact has been underestimated in traditional historical linguistics. In fact, as already mentioned, Heine and Kuteva (2005) argue that many of the common features of European languages may be accounted for by language contact rather than inheritance. A similar view had already been advocated in Haspelmath (1998), who suggests that Standard Average European (Whorf 1956) defines a linguistic area that emerged during the transition between antiquity and the Middle Ages as a result of intense language contact through migrations (1998: 285). In the study of language change, it is generally agreed that it is speakers who change languages, and not languages that change themselves (Thomason and Kaufman 1988: 4). Therefore the role of ecology should be fundamental to any area of language change. In this sense, I side with Mufwene (1998, 2003, 2008) who argues that it is the field of contact linguistics that should lead and inform our understanding of historical linguistics (and language evolution) and not the other way around. Along the lines suggested in Mufwene (2001), below I will show that external factors (here ‘ecology’) are the driving force in language evolution; moreover, I argue that, though a distinction between system-internal factors and external ecology is possible in the study of language change, this is so only at an abstract level. In reality, system-internal factors are not independent from the external ecology that motivates them (section 5.2.3.2). Therefore, the converse of contact-induced change, i.e. lack of interference, is also socially motivated. In an evolutionary view of language

change, then, contact-induced change in a linguistically diverse environment is no more, or less, ‘normal’ than language maintenance in a homogeneous linguistic environment (see also Croft 2000).<sup>5</sup>

In this respect I would like to suggest that the distinction between inheritance and borrowing, transfer etc., just as the one between normal and other types of transmission, can only be maintained in a highly marked, artificial, ideological scenario: that of a monolingual, normative, prescriptive population, typically a modern Western nation state. It is only in such a context that we can imagine the parent language involved, and it is only then that inheritance of linguistic features can be clearly assigned (i.e. calculated). On the other hand, if we look at contexts in which speakers have different languages available from which to recruit features, these distinctions become very difficult to maintain. The internal–external dichotomy in language change, as we will see in section 5.3, does not hold in an evolutionary approach, as already discussed in the uniformitarian view presented in Mufwene (2001: 139). In post-structuralist views of language and transmission, where more than one language is involved, it has been shown that the acquisition and usage of, in particular, second and third languages, is intimately related to self-representation and self-identification (Omoniyi and White 2006), and is therefore related to external, ecological factors rather than system-internal ones. This is a view of language akin to that of linguistic anthropology, where language is treated as a form of social action always situated in networks of social relations (Schieffelin 1990; Duranti 1997). The role of identity construction in language transmission will be addressed in detail in chapters 5 to 7. For the moment I simply want to open up the possibility for a view of language as a communal construct, never entirely possessed by a speaker but negotiated through speech acts between speakers (Austin 1962; Searle 1969; Mufwene 1998). In this view, genetic affiliation becomes a purely quantitative exercise, a matter of counting features, related to historical and ideological perceptions, as we can see in Thomason and Kaufman’s (1988) study of the evolution of the English language presented above.

What I hope to have argued for are the following conclusions *against* current assumptions in historical and contact linguistics:

- i. Genetic affiliation as a quantitative exercise may reveal degrees of typological relationship between input and output, but must be carried out in a comprehensive manner and properly quantified.
- ii. Lack of a communicative system in a given environment is often assumed but not proven, and there is historical, social, and logical evidence against it.
- iii. Rapidity is logically implausible and theoretically problematic; what appears abrupt is an epiphenomenon of the gaps in our historical knowledge of communicative settings.

### 4.3 Conclusions

The historical and ideological sketch of the field of historical linguistics in relation to language contact and Creole studies presented above suggests that the only ideologically plausible and empirically grounded difference between mixed language varieties and genetically less hybrid ones – assuming this can be precisely established – is an ecological one, i.e. one that relates to external factors. The ecologies in which contact language formation takes place are not exceptional in any way and have as the only common denominator a set of socio-historical traits among which multilingualism, population movements, and cultural admixture feature prominently. Though similar processes are likely to have occurred in other historical periods, it is a fact that the history of the eighteenth and nineteenth centuries in particular saw economic and political transformations which were much more intense than ever before (Sachs 2005). In this sense we witness a higher proportion of new communities emerging, some of them evolving new cultural profiles. This approach echoes calls for regarding Creoles as socio-historically – rather than structurally or developmentally – defined (Alleyne 1979; Muysken 1988, 2000; Den Besten, Muysken and Smith 1994; Mufwene 1996a, 1998, 2000; Ansaldo and Matthews 2001, 2007). Therefore, in the linguistic domain, exceptionalist scenarios that have led to the construction of the Creole paradigm, and to treating contact languages as ‘different’ in terms of genesis (e.g. Matras and Bakker 2003 on mixed languages), can be explained as being influenced by a (linguistic) ideology in which products of non-normative, non-standardized, non-monolingual ecology are regarded as exceptional. However, as argued so far, this view suffers from a Eurocentric notion of language. In this ideology, speakers of mixed codes are implicitly regarded as having gone astray in the process of ‘normal’ acquisition. But, crucially, associating normality with monolingualism is misguided; monolingual contexts are historically and culturally speaking highly marked, because monolingual, tutored acquisition is a recent phenomenon that still does not account for a majority of humanity in terms of linguistic practice. The historical, theoretical, and critical reflections presented above substantiate the following conclusions: (a) mixed codes do not have unique typological properties that set them apart from other grammars; (b) what we know as ‘normal’ transmission is actually very artificial; (c) in a linguistically realistic theory, language contact and contact-induced change are ubiquitous, and mixed grammars are normal outcomes of the linguistic ecology that defines them; (d) exceptionalist theories of contact language formation arise from the conflation of culture-specific socio-historically rooted theoretical hypotheses, not ‘objective’ scientific generalizations based on robust empirical evidence.

In the theoretical idealization developed in chapter 5, I will treat contact language formation by considering the role ecology plays in the evolution of

language. This view is in itself also a product of its time and place, and follows a growing tendency of evolutionary approaches to conceptual development (Hull 1988; Croft 2000; Mufwene 2001, 2008). The aim is to offer a framework in which ecological factors are the driving force in contact language formation, and where contact language formation is the norm that informs our generalizations on language evolution.

## Notes

- 1 Mufwene (2008) discusses the fact that Schleicher had developed his *Stammbaum* before reading Darwin, and that Darwin borrowed the cladogram used in evolutionary biology from him.
- 2 This tradition has been questioned even in monolingual communities, where variation between idiolects exists and the inputs to the learner are not identical (see Mufwene 2001, 2008).
- 3 To be fair, Thomason and Kaufman (1988: 150) acknowledge situations where substrate languages continued to be spoken for a long time, implying gradual development, as was the case for African languages in Jamaica (Singler 1984).
- 4 Whether internal and external changes can really be neatly differentiated is questionable on various grounds: Mufwene (2003: 283) claims that the changes that occur at the communal level must have occurred at the idiolectal level first; and Campbell (2002) observes that a good knowledge of the history of population movements is crucial for any successful historical linguistic exercise of reconstruction.
- 5 We will also see that both processes can be involved in the ecology of one and the same contact language, as in the case of Sri Lanka Malay (chapter 6).

## 5 Contact language formation in evolutionary theory

---

*Nothing is lost, nothing is created, everything is transformed.*  
Antoine Lavoisier, father of modern chemistry

In this chapter I propose a view of language formation and language transmission that I believe can help us better appreciate the contribution of contact language formation (CLF) to general linguistics. The field of contact linguistics has always posed a challenge for the dominant linguistic paradigm; in contact linguistics we need to overcome Saussure's dichotomy between competence and performance – i.e. language as a set of abstract features and language as usage – in order to arrive at explanations that can accommodate the role of society in language formation and transmission. In order to develop a theory of language within which the study of CLF assumes a central role, we need to look into social theories of language in order to construct a solid basis of philosophical ideas on which to build our theory.

A crucial aspect in our conceptualization of language lies in the nature of acquisition or transmission of language. In order to engage with language change over time, whether in the individual or in society, we must have an assumption about how grammar is (re)created in humans during the course of their life. How does the child obtain his linguistic knowledge? What determines the type of individual grammar (idiolect) that adults develop over time? If we want to imagine how language is 'acquired' by an individual, we could conceive of grammar as a system of rules that is passed from one speaker (e.g. the parent) to another (e.g. the child) *in toto*. This is a common assumption often found in linguistics. When this happens, we could argue, the child has successfully acquired the language. But such a view is too simplistic, it has been argued (Haugen 1972; [chapter 4](#) of this volume), because language is not a monolithic structure; it is characterized by variation and it is acquired piecemeal and selectively. So how does the 'learner' obtain the linguistic information on which to build his grammar? Even in a monolingual setting, speakers use different idiolects (see [section 5.2](#)) in acquiring language, so that a realistic – albeit idealized – picture of transmission would look like [figure 5.1](#).<sup>1</sup>

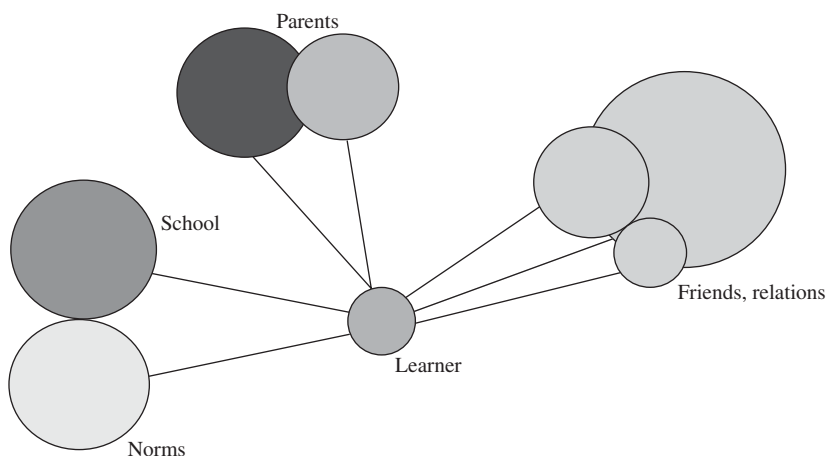


Figure 5.1 Sources of learning in a homogeneous and normative linguistic environment

As figure 5.1 shows, the learner's input languages are quite similar, but the learner certainly needs to incorporate different ways of speaking the same language, including the parental varieties, the official one, and the ones s/he is exposed to socially (all of which may also influence and be related to each other, as shown by the overlapping circles as well as the shadings). Due to the natural variation among languages, their features will be in competition, and it will be up to the learner to select certain aspects of language over others. Variation is not only phonological, which is the most obvious sign of individual differences within the same 'language'. For example, in the prevalently monolingual Italian ecology in which I grew up, though varieties mostly tended to converge to a standard form, there was variation in various aspects of grammar: personal names could be preceded by a definite article or not: *Umberto* vs. *l'Umberto*; double dative pronouns were acceptable in colloquial domains but not in school: *a me mi piace* 'I like it' vs. *a me piace* or *mi piace*. In a monolingual, normative society, educational and political institutions tend to enforce (e.g. regularize) grammatical norms to keep variation and creativity at bay by eliminating deviant features. While speakers are perfectly comfortable with variation, logically no single speaker incorporates the totality of a high number of idiolectal grammars. In the case illustrated here, the learner will not end up with seven complete idiolects, but with a selection of features from each, which s/he recombines in his/her personal idiolect, or perhaps with two varieties if s/he evolves in a diglossic society. What happens in a multilingual environment where transmission is not particularly influenced by standard norms enforced by educational and official institutions?



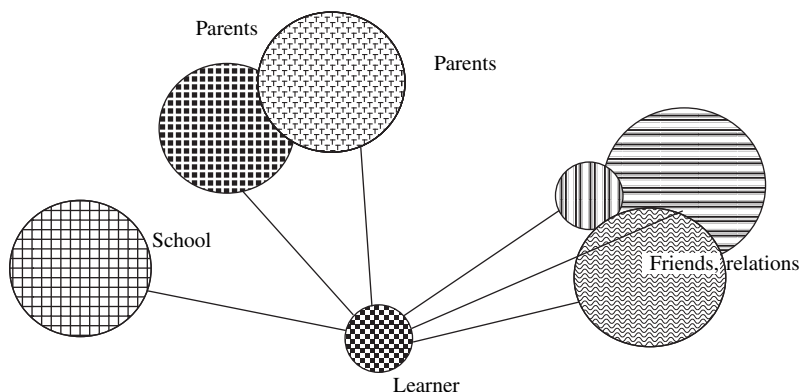


Figure 5.2 Sources of learning in a heterogeneous naturalistic linguistic environment

This is illustrated in [figure 5.2](#), where the different patterns in the different spheres reflect different languages, each of which is characterized by internal variation. This variation is not strictly limited by school or norms, though matters of prestige and typological congruence will influence the ‘weight’ of each set of features in the competition and selection process that leads the learner to become a competent communicator. What is crucial is that the learner has many more, diverse alternatives available in developing one or several distinct grammar(s). In a multilingual context, more often than not, s/he will indeed master several grammars. Thus, in a multilingual, informal ecology of transmission, the learner has an extremely rich input involving significant creativity, as s/he is less constrained by normative pressures. For example, the average Singaporean may express attributive predication in at least two ways: the Sinitic way, in which ‘adjectives’ are in fact verbal, e.g. *that book very good wat* ‘that is a good book, isn’t it?’; or with copula, as in English, also following regular English rules for tag questions. What is crucial is that, if we view transmission and acquisition as sensitive to the (dominant) varieties in which an individual grows and functions, we derive a number of observations: (a) the input is always richer than what a single output can embody; (b) there are many different possible ‘targets’ with different degrees of attractiveness depending on the social expectations on conformity; and (c) the individual always needs to select and recombine to arrive at the individual grammars that define him/her. We will discuss the implications of these observations in the remainder of this book and assess them in [chapter 9](#).

### 5.1 Competence, performance, and socialization

Not so long after Chomsky’s (1959) review of Skinner, it had already become evident that the notion of competence was too limited to account for what

really constitutes the knowledge a speaker needs in order to be really competent (Gumperz 1964, 1968; Hymes 1971, 1972). Being concerned with what is normally termed ‘performance’, these and other authors looked seriously at how to integrate variation and diversity in a theory of language. Hymes (1971), for example, replaces Chomsky’s notion of competence as tacit knowledge of grammatical rules with that of communicative competence. In this tradition, language is viewed as a communal phenomenon (Gumperz 1968), within which typically a number of varieties are represented and language is therefore regarded as heterogeneous (Hymes 1971). As a consequence of a keen interest in variation and diversity, within this tradition of thought we also find an early interest in language contact (Gumperz 1958, 1964). One common assumption in these works is that the meetings of different cultural systems seldom result in simple replacement of one for the other, but rather in the syncretic fusion of elements belonging to different systems (Hill and Hill 1986).<sup>2</sup> More recently, we have witnessed the development of a whole new paradigm centred around language as socialization; in this school of thought, language acquisition is also seen as more than mere competence, as socialization is only possible through appropriate performance (Ochs and Schieffelin 1984; Garrett and Baquedano-López 2002; chapter 7 of this volume).<sup>3</sup>

The distinction between competence and performance has also been criticized more recently in developmental studies in psychology and biology: Ingold (2000: 399) stresses how “in real life verbal composition is inseparable from performance”, and that performance is an activity that derives from contextual participation. A notion of competence as mere rules of grammar has been criticized at an even deeper level as being basically an abstraction of rules of writing, that is, a reflection of cultural knowledge which is by no means innate (Harris 1980; Coulmas 1989); the basic units of linguistic analysis such as phonemes, words, and sentences, it has been claimed, are not real mental categories but mere reflections of our experiences of reading written texts, and therefore not a universal feature of language (Olson 1994). It is important to note that these views do not question the reality of *a priori* mental categories in order to advocate nurture over nature, nor to undermine innatism in favour of determinism (Ingold 2000). Rather, these views aim to present human (linguistic) capacities as developmental properties of self-organization, neither internally predetermined nor externally imposed (see also Oyama 1985). Indeed, if we accept that multilingualism is a natural state, ‘competence’ would have to be interpreted as a combination of (a) fluency in different languages, (b) ability to code-mix, and (c) readiness to shift according to the context of communication. These are clearly practices that require socialization and thus force us to go beyond the competence/performance dichotomy.

A view of language in which performance is given more of a prominent role than (or one that is as prominent as) competence is indeed common in

anthropological linguistics, where language is seen as social action, in the sense of ‘doing things with words’, a view pioneered by philosophers of language such as Austin (1962), Searle (1969), and also argued for in Bakhtin (1981). Ochs and Schieffelin (1984, 2001) highlight the fact that the nativist position developed by Chomsky since 1959 poses at least two problems. Firstly, it relegates social and cultural observation to a peripheral role, as ‘context’, separable from actual language acquisition. Secondly, it creates an impression of the child’s learning environment as fragmented, insufficient, or otherwise lacking information for appropriate acquisition to occur. It is difficult to reconcile the view of an insufficient acquisitional scenario with the settings presented in figures 5.1 and 5.2 above; whether we consider a monolingual or a multilingual ecology, the general impression is quite the opposite, i.e. that learning occurs in an incredibly rich environment. This is particularly clear in CLF: as argued throughout this book, ignoring ecology is problematic if we want to understand the dynamics of CLF, because multilingual ecologies, in which CLF occurs, are by no means deficient in terms of input.

As already mentioned in chapter 4 (section 2), there is little empirical substance behind the idea of imperfect input, a point argued at length in Sampson (1997) who notes how the idea of ‘poverty of stimulus’ in acquisition is based on total absence of empirical observation (Sampson 1997: 38–43). A number of studies based on the observation of child–caregiver routines have shown that the speech used by adults with and around children varies across cultures and is typically well formed. Whether it contains systematic modifications of children’s speech that include high pitch, slow intonation, baby talk, and other support strategies that facilitate the learning process, or whether it is adult speech, acquisitional environments provide abundant input to the child (Ochs and Schieffelin 2001: 265). Studies in the field of language socialization have shown not only that the linguistic environment of children is rich and structured so as to facilitate acquisition, but also that children learn through direct interaction as well as indirect exposure to adults’ speech. Moreover, learning and evolution of linguistic skills does not stop in childhood but continues throughout adulthood, a fact that questions the notion of a ‘critical period’ in acquisition (Kulick and Schieffelin 2007). Indeed we can see how learning through indirect exposure and the role of adults on the evolution of language lie at the basis of the type of CLF described for the case of Sri Lanka Malay (chapter 6). Learning through indirect exposure is particularly evident in non-Western acquisitional environments, where the typical accommodation found in Western adults when speaking to children is minimal. In Kaluli (Papua New Guinea) society, small children grow up in a rich environment and learn predominantly from being spoken about, rather than being directly addressed, while Samoan children are used to adult multiparty talks from very early on (Schieffelin 1990; Ochs and Schieffelin 2001). In neither community

does simplification of adults' speech occur. Studies of language acquisition and socialization such as these ones strongly suggest that socialization routines have a deep impact on language acquisition and that language acquisition is a fundamental tool in becoming and remaining a competent member of society. This is particularly relevant to theories of pidgin formation, as we will see in [chapters 8 and 9](#), as it suggests that baby talk (and 'Foreigner Talk') may be practices specific to Western cultures and not universal dynamics of second language acquisition (Ochs and Schieffelin 2001: 272).

In language socialization approaches, language is acquired and developed through its social use by both children and adults. Socialization spans an entire lifetime and individuals appear flexible and very adaptable to change. Language can thus be seen as a set of communicative practices that speakers engage in for socialization and acculturation purposes. As Eckert (1989) argues, language is a tool used for self-representation and representation of others. These are linguistic practices that can *construct* and *transform* social structure and language can be one of the mediums, as well as a result, of the construction and transformation processes (Ochs 1996: 408). The locus of practice is the community of speakers; as the analysis of the Sri Lanka Malay communities will show, the notion of community here is not intended in the utopian sense of a linguistic community as a homogeneous, culturally discrete, and communicatively isolated group (Pratt 1987). A community has to be understood in the sense of a 'community of practice' (Eckert and McConnell-Ginet 1992; Wenger 1998), and presupposes diversity within the groups as well as shifting patterns across often overlapping communities. This implies that for an individual to be considered part of the group, a relation between membership of that group *and* interactive engagement with the group is required. A community in this sense has fuzzy boundaries and is rather permeable, in particular in the multilingual and multicultural ecologies discussed in this book. Heterogeneous contexts such as these have often displayed strong processes of identity formation, as shown in Buchholtz and Hall (2007), and it is in these typically cross-cultural contexts that social construction often leads to social transformation (Sahlin 1981). In addition, as noted in several studies of language 'shift' in multicultural ecologies, cultural elaboration and social transformation as well as active maintenance are very likely to co-occur within the same context (Field 2002). In these settings, children *and* their caregivers engage in construction and negotiation of identities, which may result in the transformation of existing codes (Garrett 2000). These processes, discussed in particular in [chapter 7](#), I will refer to as 'identity alignments' along the lines suggested in Lim and Ansaldo (2006a, 2007).

The question is: how do we integrate such views of language, as well as the ecological dimension, in a framework for the study of CLF? I believe that the answer lies in evolutionary frameworks of language and language change, in

which usage can be explained as a process of selection and replication sensitive to its ecology.

## 5.2 Language evolution and contact languages

In this section I adapt David Hull's evolutionary views of conceptual systems, as well as some of his views on science, to CLF (Hull 2001). In terms of understanding language change, this approach was pioneered in Croft (2000) and Mufwene (2001); though these two approaches differ in a number of ways, they share the basic evolutionary philosophy of language. In this view, already pre-empted above, idiolectal variation is a default aspect of language, and language can be seen as a communal construct. It is in the interaction of different lects that language change occurs in a population of speakers. Speakers select from the variables they are exposed to and 'construct' their own idiolect. Implicit in this approach is the fact that the distinction between internally and externally motivated change is artificial – all change is socially motivated – and that language develops through a combination of horizontal and vertical transmissions. In my approach, I adopt the basic principles of the 'Theory of Utterance Selection' (Croft 2000), which builds on Hull's (1988) generalized theory of conceptual evolution, and combine these with the ecological approach of Mufwene (2001) in terms of the role of external and internal ecologies in language evolution.

In addition to an evolutionary approach, I present a functional-typological and usage-based approach to language (following Bybee 1998, 2006; Croft 2000, 2006a; Tomasello 2003).<sup>4</sup> The rationale behind my approach to CLF rests on three related ideas:

1. As already mentioned in chapter 4, it is widely acknowledged that it is speakers who change languages (Thomason and Kaufman 1988). Therefore, if we want to understand the process of CLF, we must integrate grammatical analysis with sociolinguistic theory.
2. Whichever theory of language and grammar we subscribe to, language involves at least a set of general cognitive strategies and some core strategies of transmission that all humans have in common (e.g. Tomasello 2003).<sup>5</sup>
3. Multilingual ecologies and informal transmission are 'normal'; they differ from monolingual ecologies in that they offer a broader range of variation (Mufwene 2002b).

Following from the three points above, this framework focusses on CLF as a communal and not an individual process, and it assumes that the patterns of language usage and language change observed in CLF must contribute to general explanatory frameworks and do not require exceptional scenarios.

### 5.2.1 *Evolutionary frameworks in linguistics*

The first in modern historical linguistics to present an evolutionary view of “language as a population of variants moving through time and subject to selection” was most likely Roger Lass (Lass 1997: 377).<sup>6</sup> Though the absence of the role of speakers in language change in Lass’s proposal justifies a certain scepticism (Milroy 1999; Janda and Joseph 2003), I think that it is important to appreciate the fundamental positive intuition in the work that is consistent with current notions of evolutionary theory. Lass (1997) expresses a belief in a theory of a ‘historically evolved system’, in which both biology and language converge. In Lass’s view, evolutionary theory may help us explain language change, not because languages are seen as biological systems, nor because language is seen as a genetic feature in the Chomskyan sense. Lass’s idea is that linguistics and biology may be seen as systems that are in some ways similar, most notably in the fact that they evolve over time and involve modification. The most important question we should worry about, as he aptly puts it, is the non-essentialist one:

Which (kind of) facet is the best one to look at given the particular epistemic game one happens to be playing? From this point it doesn’t really matter if one is a realist or an instrumentalist, if one wants to grant priority to one particular facet or another. *The point always is fruitfulness, not ‘truth’*. We are Model Builders, we are not (or should not be arrogant enough to think that we are) in the Truth Business. (Lass 1997: 384–5; my emphasis)

Lass was partly inspired by Dawkins (1976), who suggested that his work on evolutionary biology could be applied to the evolution of cultural systems. The idea therefore is not that language, change, and contact are the same as biological evolution, but rather that they share interesting common properties worth exploring, a view embraced and expanded by Hull (1988, 2001), Croft (2000), and Mufwene (2001, 2008). In a critical take on evolutionary theory in linguistics, Andersen (2006) highlights a number of mismatches between analogies drawn in the fields of language and biological evolution. This however is a common reaction that fails to understand (a) that it is not necessary to attain a total overlap between explanatory models in order for the analogies to be useful (Croft 2000, 2006a, b), and (b) that, at least in the sense of Croft (2000), it is the field of *conceptual evolution* – and not biological evolution – that provides the necessary foundations for a linguistic framework (Hull 1988). Moreover, Andersen reveals a strong belief in typical products of European historical linguistic scholarship, such as the *Stammbaum* model, already discussed in section 4.1. Therefore, the fact that there may be mismatches between evolutionary patterns and so-called mechanisms of change should rather be seen as a productive outcome of applying a novel framework to established concepts, as it reveals their inherent limitations.

There are thus two important views expressed here: the view of language as a historical entity, and the role of the speaker in language evolution. Both of these views fit perfectly in a functionally based evolutionary take on language. Moreover, from a functionalist perspective, language (and grammar) *is* a historical entity shaped by speakers' use. It is useful to embrace this if we want to tackle the dimensions of variation and diversity that characterize language use, particularly in multilingual, heteroglossic environments typical of contact languages (section 5.3).

### 5.2.2 *A theory of selection*

In order to clarify the basic notions of evolutionary theory applied to language change, I briefly revisit some of Hull's basic ideas. Before we start, it is useful to remind ourselves that there is no unanimous agreement among biologists regarding the notions of 'ecology', 'inheritance', and 'selection'. The interested reader can find a discussion of some current trends in Hull, Langman and Glenn (2001). It is also necessary to clarify that any conceptual analysis can engage at different levels of enquiry and that these levels may require different principles. As Hull (2001: 58) points out, it is difficult to know whether one is at the 'right' level of analysis; it is possible that a more general, or a more specific level, may, in a given instance, yield better results. Different levels are all open to investigation and equally valid in linguistics alike: the individual speaker, the household, the network, the wider community, all offer interesting and useful perspectives on the formation of contact languages. I argue below that the main issue in linguistic analysis is to not confuse the levels and the principles that derive from them, and to justify the relation between the domain of enquiry and the theoretical realm. Finally, evolutionary theory is a theory of *change*; as such it is first and foremost concerned with understanding the past, not with predicting the future. Since CLF involves the study of linguistic changes that occur in the formation of new grammars, it is not a far-fetched metaphor to look into evolutionary theory for a better understanding of the relevant processes. Rather, it appears to me to be a logical step in the development of science.

Hull (2001) warns us against brief definitions of selection.<sup>7</sup> In order to express what selection can entail, he formulates it as follows (2001: 53): "selected cycles of replication, variation and environmental interaction so structured that environmental interaction causes replication to be differential". This definition in itself is not enough, and must be elaborated with explanations of variation, replication, and environmental interaction as outlined in the following points.

- Variation: Variation can be considered as a precondition to the selection process (Hull 1988, 2001). Variation is absolutely essential in selection; without



it there are no alternatives available for selection. Variation is totally caused and not blind or random, contrary to what is often claimed in criticisms of evolutionary theory outside biology. What causes it is what we should aim to find out (Hull 2001: 54). With respect to the adaptive function of variation, whether or not these are neutral to the environment can be debated (see Hull 2001: 55 for specific discussion and further readings).

- Replication: Replication is a copying process. In Thompson's (1994) view, natural selection fundamentally requires that the presence of certain patterns of elements in one generation make the presence of those same patterns in the next generation more likely. In this sense DNA inheritance is not required, though most current approaches to evolution assume it, but there is no consensus about what essentially is inherited (Hull 2001: 59).
- Environment: Environmental interaction causes differential replication through selection. Differential replication is just as important as mutation (or replication over time) in evolution. The environment can be understood in terms of ecology.

In short, variation, replication, and environment are the building blocks for a theory of evolution. In addition, a note on the actual locus of change is required: Hull (1988) assumes a population theory of species. In a population theory, a population is defined as a set of interbreeding individuals who are reproductively isolated from other populations; it is within the species that change takes place. This is a departure from an essentialist view of species, where a species is identified on the basis of a set of constant structural properties. A population is spatio-temporally bound; there are no inherent properties that define the set, but rather relational properties between members of the set. It is important to note that it is only through sexual reproduction that new species can be formed. With these basic notions in mind, let us see how they can be applied to language change, following Croft (2000, 2006a, b).

### 5.2.3 *Language change in evolutionary perspective*

In Croft (2006a), language change is simply seen as the consequence of *linguistic variation* at a broader level; his model distinguishes between three different orders of variation. First-order variants are individual variants that occur in all subdomains of grammar (phonological, semantic, and syntactic). When first-order variants take on a sociological significance in a community, they become second-order variants. When second-order variation becomes conventionalized in a speech community and leads to divergence, we have third-order variation (i.e. variation across languages).

#### 5.2.3.1 *The theory of utterance selection (TUS)*

Based on Hull's (1988) work on the role of replication in evolution, Croft (2000) proposes the following generalizations (my adaptation). As we can



see, points (a) to (e) derive directly from Hull's generalizations presented in the previous sections as applications of his notions of change, replication, and population to language:

- a. In language, as in other complex evolving systems, change is caused by *replication*.
- b. In communicating, speakers replicate linguistic features they have been exposed to. This replication can be more, or less, faithful, resulting in 'normal' or 'altered' replication. In what follows I use the terms 'identical' and 'innovative' instead.<sup>8</sup>
- c. Change happens at the population level. A population is *a set of linguistically interbreeding individuals*. The locus of language change is therefore a speech community or a social network, where identifiable (Croft 2000: 17–20). A population can split into 'varieties' because of (relative) geographical isolation; this allows us to conceptualize a phylogenetic classification of languages.
- d. Specifically, a language is interpreted as a population of utterances in a speech community; a population view of language suggests that a speech community is not defined by an abstract linguistic system (an essentialist position), but rather by cultural and socio-historical factors.
- e. Grammars are idiolects: each speaker's knowledge may include more than one grammar.

Simply put, Croft proposes that, in communication, speakers engage in the following activities: they select (features of) utterances for replication; they replicate, and in so doing they might innovate (2000: 38). Finally, they are involved in the propagation of replications within the speech community (2000: 4). His view of language change is summarized in the theory of utterance selection (Croft 2000: 39):

- TUS does not preclude other selection processes occurring at other levels of language.
- TUS does not require a specific set of causal mechanisms for the selection of utterances.
- TUS puts linguistic convention at the centre of language change: normal replication means conforming to conventions of the speech community (i.e. it has socio-historical motivations); altered replication means departure from those conventions, i.e. innovation.

From the point of view of TUS, utterances matter and not sentences, meaning that it is the *actual occurring linguistic features* (texts, dialogues, etc.) that are relevant. This is obviously a theoretical stance on what constitutes a grammar and language. The idea is to distance the model from an essentialist notion of language (just as the population view of species distances itself from an

essentialist view of species), because a structuralist definition of language has the same problems as the essentialist view of population, turning a historical entity into a type (see [section 4.1](#) and Ansaldo 2007). If a population is a set of spatio-temporally-bound actual individuals – a set of ‘possible’ individuals makes no sense – a language should be defined by the actually occurring phrases, not the ‘possible’ ones. This point relates to the notion of frequency of usage discussed in [section 5.3](#) below.

### 5.2.3.2 *Selection, replication, and propagation in contact language formation*

In what follows, I move on to apply all of the above directly to the evolution of new grammars in CLF. As argued above, in an evolutionary view of CLF, speakers in a multilingual community are exposed to substantial variation, i.e. they have a very heterogeneous set of utterances from which to select material for replication. As already suggested in the seminal work on contact by Weinreich (1953), I assume that an extremely low level of multilingual competence is sufficient for different linguistic systems to interfere in a speaker’s use of language, and that new language formation is the natural, creative process that typifies CLF (also Heine and Kuteva 2005).

In an evolutionary view then, CLF is defined as involving three processes:

- i. Selection of variants from a highly diverse and large pool of features.
- ii. Innovative replication and identical replication.
- iii. Propagation of variables in a population.

As we can see, these three processes reflect Hull’s (2001) definition articulated in terms of variation, replication, and environment. In order to properly understand this, we need to investigate (a) what causes variation, (b) how replication occurs (the patterns of copying), and (c) what is the nature of the ecologies that define CLF.

- (a) Variation and selection, I will argue ([chapters 6 to 8](#)), are caused by social and demographic dynamics. There is plenty of literature on this topic, from variational sociolinguistics to the study of ethnolinguistic vitality. As we can see in the formation of Singapore English, notwithstanding the fact that English is a powerful presence in the lexicon due to its economic capital and institutional support, different linguistic features selected at different stages of evolution tend to reflect the dominant language(s) of each phase (Lim 2007). The features of Malay selected in the early ecology of Singapore English can be seen as founder effects that survive to this day (Lim 2007, 2008, 2009a).

- (b) How linguistic variables interact (the ‘competition’, if you like) depends largely on their typological and functional properties (Croft 1995); this can be seen as the internal ecology discussed in Mufwene (2001) as the structural make-up of the environment in which CLF takes place. This phase can thus – to some extent – be investigated at the system-internal level, though it is never completely independent from socio-historical principles. As I will argue in [section 5.3](#), different types of frequency need to be identified as a central factor behind the evolution of grammar. For example, in accounting for the emergence of case in Sri Lanka Malay, we can see how token-frequency, i.e. frequency of occurrence, and type-frequency, i.e. frequency of common structure (congruence), together determine the structural output of CLF (Aboh and Ansaldo 2007; Ansaldo 2008, 2009a; [chapter 6](#) of this volume).
- (c) Propagation needs to be distinguished from selection. Propagation really means the diffusion of the new patterns, i.e. the stabilization of grammar. Here the role of ecology is fundamental, as propagation relates to the type of environment in which selection and replication take place. This requires considering the type of community that propagation occurs in, whether close, open, diffuse, stable, etc., and the size of community, i.e. large or small. For example, size determines rate of change: logically, propagation will be relatively fast in small and tightly knit groups, but must be slower in large and diffuse communities. As shown in Ansaldo and Nordhoff (2009), morphological case markers developed in Sri Lanka Malay in roughly 300 years of history. In other contexts, e.g. Old Persian, a similar process is known to have taken approximately 600 years to come full circle. Does this mean that case in Sri Lanka Malay evolves rapidly? No, the authors argue: it simply means that it takes less time for such an innovation to stabilize in a community of fewer than 50,000 individuals than for the same stabilization to occur in a much wider and geographically more diffuse population. This explains why CLF, most typical of often small and tightly knit groups, has been often described as ‘rapid’ in the literature (see also Matras and Bakker 2003; cf. Ansaldo and Nordhoff 2009). This does not mean that small groups necessarily undergo frequent change. It simply means that small groups efficiently manage usage. In this sense, in a different ecology a small group may just as well be expected to display very strong maintenance, as argued in Nettle (1999) and Trudgill (2002) for adoption and maintenance of marked features.<sup>9</sup>

Note that the three ‘processes’ listed above are not really separate phases with a hierarchical ordering, as clearly argued for evolving systems in general in Hull (1988, 2001): selection, propagation, and innovation may occur simultaneously and iteratively at any level of organization.

### 5.2.3.3 *The ecology of CLF*

Ecology determines the output of CLF thanks to its external and internal factors. External factors relate to socio-historical and cultural features of the contact environment; an understanding of the external ecology of Monsoon Asia, presented in [chapters 2 and 3](#), allows us to determine the salient sociolinguistic dynamics as well as the varieties involved in the CLF processes that characterize the region. Ecologies can also be internal to species (Mufwene 2001: 22–3) in the sense of the variation and interdependences that obtain among features in a language, especially how each feature is influenced by other variants and/or other features in the overall system. Looking at variables from different languages as participating in a pool from which they can be selected and replicated (Croft 2000) is particularly useful for a typological approach as it offers the investigator a well-defined domain in which to study areas of (non-)congruence between the different grammars in contact.

The notion of feature pool (Mufwene 2001) allows us a holistic picture of CLF processes in which the structural features of the varieties in contact can be investigated, as shown, for example, in Aboh and Ansaldo (2007) for Suriname Creoles and Sri Lanka Malay. Along similar lines, in Ansaldo (2004b) I introduced the notion of typological matrix (TM) as a parallel concept to that of feature pool. In the remainder of this book I will keep using the term TM when discussing the internal ecology of CLF, i.e. system-internal considerations. It is important to note that I do not suggest that we consider internal and external ecologies as separate dimensions; since the external ecology represents the socio-cultural and historical dimensions of language, and since language in my view is a socio-cultural and historical phenomenon, it follows that the internal ecology – or TM – is actually a complement of the external ecology, as illustrated in [figure 5.3](#).

The notion of TM offers a convenient abstraction for the purpose of function–form analysis, and allows us to distinguish between socio-historical investigation, which is required to understand selection and propagation, and actual innovation, i.e. different patterns of recombination of variables. Nonetheless, as we will see, external and internal factors interact with each other in a complementary way and cannot always be clearly separated (see [chapter 6](#)).

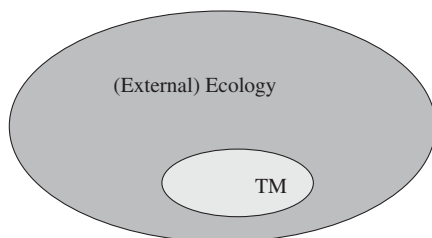


Figure 5.3 Internal and external ecology

The notion of TM does not differ significantly from the notion of feature pool put forward in Mufwene (2001). Mufwene's notion implies a distinction between internal and external ecology, and identifies salience and congruence, among others, as significant factors determining the outcome of the selection process. I am in agreement with these notions, and TM can therefore be seen as interchangeable with the notion of feature pool in the sense of internal ecology mentioned above. The difference is that, in the TM, I try to reinterpret notions of markedness, congruence, and salience in terms of type- and token-frequency patterns, along the lines put forward in Haspelmath (2006), as explained in [section 5.3](#) below. The notion of TM suggests that, in CLF where variables of different typological origins interact, it is in functional-typological analysis that we find the key to understanding *how* new grammar emerges. This is particularly relevant in the context of this volume as a typologically based approach predicts that the output of a contact ecology is heavily dependent on the input.

Thinking in terms of TM is really a methodological issue. A good example is provided by the study of the evolution of Singapore English or Singlish, formerly regarded as a 'creoloid' (i.e. a Creole without a pidgin ancestor, Platt 1975), more recently as a new variety of English (Pakir 1991; Schneider 2003). In looking at the grammar of Singlish, for example, many scholars have taken a contrastive approach, typically comparing the contact variety to the putative lexifier, usually taking the standard as the point of departure, in our case Standard English. But this approach has obvious limitations because in CLF no single language constitutes the exclusive target ([chapters 6 and 7](#)); rather, patterns of code-mixing involving different typological profiles account for much of the communication process. In the Singapore ecology, for example, for the past couple of centuries, code-mixing involving the following languages has been part of the communicative norm: (a) English vernacular and Standard English (at present), (b) Bazaar Malay, (c) Min varieties (Hokkien and Teochew, and (d) Cantonese (and other Sinitic varieties). Today, with a somewhat more stable Singlish functioning as one of the two 'native' languages of a majority of Singaporeans (the other being Singapore Mandarin), the Singapore ecology can be characterized by selection and replication processes from a TM dominated by Singlish, Standard English and Singapore Mandarin, with Hokkien, Cantonese, and other languages still featuring to some extent (Lim 2007, forthcoming). In a TM approach, any grammatical investigation of Singlish should take these typological profiles into account (see Ansaldo 2004b, 2009b).<sup>10</sup> Similarly, as I will show in [chapter 6](#), a methodologically sound investigation of Sri Lanka Malay cannot ignore the interplay of the three dominant varieties in the matrix, namely Sinhala, Tamil, and Malay.

At the outset of this section I defined an evolutionary approach to language as fundamentally relying on the notions of variation and selection. The rest of

the section was dedicated to explaining how these notions find their place in a theory of CLF. One aspect that still needs to be addressed is the implication that speakers do not just inherit linguistic traits but actively select them from the environment to construct their idiolect. This idea is corroborated by the theories presented in [sections 5.3 and 5.4](#).

### 5.3      **Functionalism assets for contact linguistics**

In functionalist and cognitive theories of grammar, linguists tend not to isolate the structure of language from language use, as grammar is intended as the cognitive organization of a speaker's experience with language (e.g. Givón 1979a; Langacker 1987). Grammar in both cognitive and functionalist theories is seen as a set of cognitive representations that rely on general abilities of categorization and generalization, and are therefore not specific to linguistic competence. While this view has until recently not been the dominant one in linguistic theory, as opposed to the orthodox, structuralist view of grammar developed in Chomskyan linguistics, in recent years a number of trends have been accumulating substantial evidence in favour of what is clearly a usage-based theory of language, in particular the field of grammaticalization studies, language acquisition, and functional-typological linguistics (Bybee 1998, 2006). A very important discovery that these approaches share regards how living languages develop new grammar. Language, according to Bybee (2006: 730), “can be viewed as a complex system in which the processes that occur in individual usage events ... with high levels of repetition, not only lead to the establishment of a system within the individual, but also lead to the creation of grammar, its change and its maintenance within a speech community”. Research in these areas over the past two decades shows that grammar evolves naturally out of pre-existing lexical material in language use, as is abundantly documented in the literature on grammaticalization (Hopper and Traugott 1993). Grammar is a constantly evolving system in this view and “there is every reason to believe that all existing grammar came about in just the same way we observe in the documented cases at our disposal ...” (Bybee 1998: 250). Grammar therefore can be seen as an emergent system, complex and dynamic, similar to other complex systems observed in physics and biology (Lindblom, McNeilage and Studdert-Kennedy 1984; Hopper 1987). Such views of language offer powerful insights for our understanding of CLF as they incorporate what I have identified as the most fundamental factors in CLF: variation, replication over time, and speakers' agency in evolution.

#### 5.3.1      *Usage and transmission*

Probably the most systematic evidence to date in favour of a usage-based view of language can be found in the work of Tomasello, in which empirical

psycholinguistic evidence and theoretical engagement lead to a definition of linguistic competence that recognizes three phases (Tomasello 2003): (i) phylogenetic, as the biological adaptations enabling members of a species to communicate linguistically; (ii) historical, i.e. the process through which a set of conventions emerge in a given speech community; and (iii) ontogenetic, in the sense of the development that leads an individual to acquire full competence during life. In accordance with much work in cognitive-functional grammar (e.g. Bybee 1985; Langacker 1987; Givón 1995; Croft 2001), patterns of usage are ritualized (or grammaticalized) into constructions, and new generations inherit such constructions through exposure. The ability of abstracting general grammatical rules, Tomasello shows, is enabled by biological cognitive skills that are not specific to language, but fall under the categories of ‘intention-reading’ and ‘pattern-finding’. Grammar in this view is therefore not a biological adaptation but a historical and ontogenetic process. Crucial to this view is the notion of dual inheritance, which in biology emphasizes that organisms inherit both their genes and their environments (e.g. Durham 1991). This applies to humans as well, but with a twist, as humans are adapted both for pre-existing structures in their environments as well as for acquiring completely new knowledge from the socio-cultural context (Tomasello, Kruger and Ratner 1993). Therefore humans acquire language because they are biologically prepared with cognitive strategies that among other things allow linguistic categorization, and because they are exposed to their own linguistic culture.

### 5.3.2 *The role of frequency*

The above-mentioned approaches to language all share the basic assumption that structural aspects of language need to be explained by making reference to linguistic function, or usage. This is because in usage-based linguistics, as argued above, grammar is seen as a dynamic system, which varies within the lifespan of adult speakers and in which phenomena of language use do influence the representations of grammatical knowledge. In particular, there is wide agreement that frequency plays a dominant role (Bybee 1998, 2001, 2006; Croft 2000, 2001; Langacker 2000). Frequency of usage, together with structural similarity, is responsible for the depth of entrenchment and the degree of abstract generalization of grammatical rules (Bybee 2006; Croft 2006a; Haspelmath 2006). In recent quantitative analysis of historical linguistic corpora, Pagel, Atkinson and Meade (2007) confirm that frequency of usage guarantees selection over time in the evolution of Indo-European languages. Frequency is also responsible for regularization, as shown by Lieberman *et al.* (2007) for the evolution of irregular verbs in English, where irregular verbs of low frequency regularize over time while high frequency ones do not. As noted in Fitch (2007), frequency effects can account for at least 50 per cent of the variation examined in these two studies, a level of statistical accuracy that is

uncommon in comparative studies in the social sciences. As we will see below (and in [chapter 6](#)), frequency is a fundamental factor behind contact-induced change in an evolutionary framework, as it allows us to understand much of what gets replicated and how the replication happens.

As early as 1966, Greenberg had already commented on frequency as a fundamental factor in the emergence of grammatical categories and for our understanding of synchronic states (Greenberg 1966). More recently, Bybee (2006) offers a review of frequency effects relevant to the processing and storage of grammar. The main claims are the following.

1. High-frequency words and phrases undergo phonetic reduction at a faster rate than low- and mid-frequency sequences, e.g. *do not* > *don't*. This can be explained by the fact that articulatory representation depends on neuro-motor routines, which increase in fluency when repeated. The increase in fluency can lead to a group of words being identified as a single unit.
2. High-frequency strings become more entrenched in their morphosyntactic structure and more resistant to restructuring. Frequency strengthens the memory representation of words, phrases, and syntactic structures alike, making them easier to access and more conservative (Bybee and Thompson 1997). Bybee (2006: 715) offers the example of high-frequency irregular verbs in English such as *keep*, *kept* that maintain their irregularity while low-frequency ones such as *weep*, *wept* are more prone to regularization (*weeped*).
3. Related to the above, high-frequency strings can lose semantic and syntactic transparency and become autonomous (e.g. words with derivational affixes that increase in frequency lose the transparent relation to their base form, Bybee 1985).

Based on exemplar theory developed, in particular, for the representation of phonetic variation (Pierrehumbert 2001, 2002), Bybee (2006, 2007) offers the following observations on the relation between use and cognitive organization:

- (i) Organization, or storage in memory, is probabilistic, not categorical. While an adult who experiences a new string will be only minimally influenced as s/he already possesses a wealth of older, more entrenched strings, a child for whom most items are new will be more influenced in its organization. Repetition or ritualization increases ease of access (Haiman 1994).
- (ii) Repeated linguistic experiences, just as experiences of other kinds, lead to strong memory of such experiences and loss of non-repeated ones (Goldinger 1996).
- (iii) Linguistic memories do undergo reorganization.



Frequency effects are thus likely to play a significant role in the evolution of grammar, independently of the ecological environment in which such evolution takes place. In particular, we can talk about two broad types of frequency: (i) token-frequency, i.e. the frequency of a linguistic unit in text or discourse; and (ii) type-frequency, i.e. the frequency of a pattern (or construction) in grammar(s).

In analyzing the TM of a given contact ecology, it is therefore important to understand, when possible, the different frequency patterns that occur (Croft 2000, 2006a). In the process of CLF, token-frequency can be interpreted as discourse frequency: linguistic items that are frequent in discourse may be those that are grammatically obligatory, semantically salient, or pragmatically more relevant (Aboh and Ansaldo 2007; Ansaldo 2008, 2009a). Type-frequency can be interpreted in terms of congruence: constructions that are more common, for example, because of typological congruence, are more likely to dominate the TM. For example, in Sri Lanka Malay, a case system evolves as a result of ecological factors. Case marking is a typological trait of the two significant adstrates in the TM of Sri Lanka Malay, namely Sinhala and Tamil. Because of congruence, or type-frequency, Sri Lanka Malay also develops a case system, even though the third element in the matrix, the Malay ‘lexifier’, lacks case marking altogether. In particular, the Dative case marker, which has very high token- (or discourse-) frequency, as well as high type-frequency – there is full congruence between the two adstrates – emerges in Sri Lanka Malay identical in function to the adstrate patterns. The Accusative case marker, on the other hand, with low token- and type-frequency – not obligatory and divergent in the two adstrates – evolves as an innovative replication of selective features of the two adstrates. This illustrates how differential frequency patterns in the TM lead to different outputs within the same ecology (see chapter 6). Another example is found in zero-copula structures in Singlish: these are not the result of simplification, reduction, or imperfect acquisition of English. Rather, in the TM in which Singlish evolves, zero copula is found in all adstrates, be they Sinitic or Malay varieties: this means that type-frequency is high and token-frequency can be inferred as a consequence of type-frequency in this particular ecology. Therefore zero copula is the dominant feature and thus emerges in Singlish grammar as the preferred strategy (Ansaldo 2009b; see also chapter 7).

Note that frequency in this sense, as argued in Haspelmath (2006), can help us make sense of the different interpretations of the notion of markedness. Typological markedness has been used variably to refer to (1) cross-linguistic rarity (Jakobson 1963), (2) conceptual difficulty (Givón 1991), and (3) deviation from parametric options (Chomsky 1981). All these interpretations can be traced back to observations of cross-linguistic frequency. Cross-linguistic frequency however is not an objective indication of what are more or less ‘natural’

linguistic patterns but rather a result of sampling methods, linguistic history, and theoretical frameworks (Croft 1990; Newmeyer 2000; Haspelmath 2006). The notion of conceptual difficulty – or cognitive complexity – is likewise related to what is really structural familiarity. It is indeed difficult to qualify complexity beyond the language-specific domain; Ansaldo and Matthews 2001, for example, criticize the notion of tone as a complex feature of language as implied in the Creole Prototype (McWhorter 1998), noting that, to speakers of tonal languages, tones are surely not complex. In point of fact, Lim (2007, 2009a) finds that tonal features are present in the grammar of Singlish, as a consequence of the prevalence of tone within the TM in which Singlish evolves. In a parallel case, as we will see in chapters 7 and 8, it is doubtful whether zero copula can be considered an unmarked option: it is surely not less complex than overt copula, as it requires attributive adjectival semantics, arguably a more complex semantic class.<sup>11</sup> As a solution to the cross-linguistic relativity of the prevalent usages of markedness, in recent work I have opted for the notion of ‘local’ markedness, i.e. relative to a specific feature pool (Aboh and Ansaldo 2007; see also Mufwene 2001 regarding ‘ecology-specific markedness’). Frequency (or markedness) may not explain everything in CLF, as shown in the case of Sri Lanka Malay where founder effects also determine the evolutionary course; what I hope to contribute with this approach is to find out just how far frequency patterns can help us explain the input–output relation in the study of CLF.

## 5.4 Conclusions

Tutored, normative acquisition in a monolingual setting is not the norm in the history of the world, where multilingual practices and informal transmission of language have by far been prevalent in many of the world’s societies (cf. Mufwene 1990; DeGraff 1999). It is problematic that the former has been taken as the default in the formulation of our theories of language. I have suggested above that this is the result of a confluence of ideological traits of the historical period in which the study of historical linguistics and language change emerged. As we have seen, a combination of replication and selection can be taken to underlie evolutionary processes in general (Hull 1988, 2001), and this can be successfully applied to CLF in particular. Following from this and the previous chapter, it becomes difficult to draw a clear line between ‘pure’ and mixed languages, internally vs. externally motivated change, inheritance, and borrowing. Research in typology and history of Creoles clearly reveals the complex relationships these languages have to their sub- and superstrate languages (see Ansaldo, Matthews and Lim 2007). Hjelmslev (1938) already pointed out that all languages are mixed to some extent. Mufwene (2001, 2005, 2008) proposes a ‘uniformitarian principle’, which essentially accords all languages contact-based evolutionary paths.

In accordance with the theories illustrated in this chapter, languages are best understood as constantly evolving systems that defy static taxonomic categorization. Along the lines proposed in Mufwene (2001, 2008), I have suggested that we view language as a population of idiolects that constantly evolves (see also Mufwene's characterization of 'communal language' as a 'species'). Following Croft (2000), the locus of variation is communication, i.e. the interbreeding of different idiolects. This means that when speakers interact they exchange utterances; in such an exchange, they replicate linguistic features. Speakers may replicate the linguistic features of their environment (their community, network, tribe, etc.) identically, or they may replicate them with innovations. I propose that identical replication is what we might ideally expect more often than not in an environment with a high degree of monolingualism and strong normative tendencies (e.g. a clear notion of standard and compulsory, institutionalized education). In order to understand the types of selections that occur in CLF, we need to focus on the dynamics of the external ecology, as well as on the functional alignments of the internal ecology. (Note that even in such environments where linguistic diversity is limited, language change does still occur, as it does in non-contact environments, because speciation also occurs in relatively isolated communities, which means that innovative replication happens in language transmission between speakers in all kinds of environments.) It is in a linguistically diverse environment with low normative tendencies that innovative replication is most likely. CLF is thus the cumulative outcome of the innovations in a population of speakers.

This chapter has introduced the tools and the concepts required in order to deal with the inherent variation in language that leads to speciation, in particular the diversity of features in multilingual transmission and the linguistic practices of speakers in these environments. In [chapter 6](#), I will illustrate in detail how to reconstruct the external ecology of a CLF environment; I will also focus on the functional alignments within the matrix, in order to illustrate the role of frequency in the restructuring process. [Chapter 7](#) will focus, in particular, on typological admixture as the manifestation of selection and replication processes. In [chapter 8](#), I will discuss the role of different ecologies in language evolution, and address the issue of pidgin versus other ecologies.

Finally, it should be noted that the framework for CLF and the notion of TM presented here are, as all evolutionary explanations, primarily tools to understand the past; whether they may help us predict future developments or not is only a secondary question, as this is not what they are developed to do. Structuralist linguistics is influenced by the positivist belief that for a theory to be worthy it must have predictive power. This actually derives from a rather superficial understanding of the philosophy of science. Hull (2001) makes it clear how epistemologically novel predictions are actually less significant than they appear: a theory that is sufficiently well constructed will indeed imply the

occurrence of a certain phenomenon that, upon investigation, will be ‘discovered’ by the scientist. What is observed has been derived by the theory, so its epistemological status is unclear. Moreover, the whole idea of testing a hypothesis rests on the belief that, this way, ad hoc theories will be avoided. But in the progress of science, Hull explains, hazy theories and scanty data are a reality and cannot always be avoided. Against this positivist school of thought, the historical school has claimed that when old facts are explained in new ways, these count as new facts (Lakatos 1970). In this line of research, scientific methodology must be historical and is therefore necessarily retrospective; this is in tune with the evolutionary theory presented here.

## Notes

- 1 The spaces in between the different spheres in [figures 5.1](#) and [5.2](#) are for clarity of illustration only; in reality all spheres within a given ecology overlap to some extent.
- 2 This is not to say that all multilingual communities necessarily produce contact languages, but that language admixture is a common occurrence in such environments.
- 3 Duranti (2001) notes that, from these perspectives, Chomsky’s notion of a unitary language is not just a working hypothesis but rather an ideological stance.
- 4 These are theoretical and methodological ‘choices’ that determine the primary domains of investigation and that may vary depending on the theoretical stance of different approaches. I do, however, point out, when necessary, how this theoretical-methodological orientation provides some advantages over competing alternatives.
- 5 Though cross-cultural differences in linguistic routines are acknowledged, these differences can be seen as matters of frequency and/or contexts of occurrence (see Ochs 1986: 10).
- 6 Though the idea is not entirely new: Keller (1994), for example, suggests that we view language as an ‘object of the third kind’, i.e. not a natural product, not a human construct, but the product of a conceptual evolution.
- 7 Hull (1988, 2001) is first and foremost interested in resolving a number of controversies in evolutionary biology, specifically regarding the objects of selection in the neo-Darwinian debate. At the same time, he intends to offer a view of selection that transcends biology and can be applied to conceptual systems at large. For the purposes of this study, and for a general framework of language contact in evolutionary terms, it is not necessary to grasp the actual debate which gives rise to Hull’s view, but rather to highlight the product of his work and how it relates to language contact and language change.
- 8 It can be argued that replication is never really identical, as no two individuals use the same pronunciation of a word, and grammatical variation is inherent in any idiolect.
- 9 We will also see that the type of ecology determines the stability of a grammar, as in the case of China Coast Pidgin ([chapter 8](#)).
- 10 The TM can be a useful tool in two different ways: as we will see below, in cases where we do have substantial information about the ecology of a CLF process, the

logical way to proceed is from the analysis of the ecology to determining the TM; it is ecology that triggers and determines replication and it is the social history of speakers that determines language change. In cases where the relevant ecology has changed and is no longer observable, but we do have the output of CLF, the analysis of the TM may help us reconstruct a plausible ecological environment and add therefore to our understanding of the triggers and processes that define the evolution of a new grammar.

- 11 I do not discuss parametric settings in this context, as I do not engage with a parametric view of grammar (see Newmeyer 2005 for a discussion of the relationship between typology and parameters).

## 6 Congruence and frequency in Sri Lanka Malay

---

*The wives, which in part are Amboinese, in part Singulayans, and Malabarians, say nothing against this, but when the men game away their little property, they must nourish him and his children as well as they can through the month and await his better fortune at gaming.*<sup>1</sup>

Christoph Schweitzer (ca. 1630)<sup>2</sup>

The story of the ‘Malays’ of Sri Lanka and their language is one of the most interesting and instructive examples of the formation of a new language as the result of language contact and identity alignment. Note that ‘Malay’ is a misnomer here: the term was used during British rule to classify people who came from Java and other parts of the Indonesian archipelago. The Dutch term ‘Javaans’ would be more appropriate for the ancestors of the Sri Lanka Malay (SLM) people, who, like other Malay Diasporas (e.g. the Cocos Malay), indeed refer to their own language as *Java* (see [section 2.2.3](#)). Unlike its better-known Caribbean Creole counterparts, SLM, like a few other varieties of the region (e.g. Baba Malay, Cocos Malay, see [section 3.2](#) and [chapter 7](#)), is typologically in a unique position of providing us with an environment in which no (Standard Average) European acrolectal variety is involved in the dynamics of contact. Furthermore, with Sinhala and Tamil, its adstrates, the languages involved in the formation of SLM varieties come from three distinct language families: Austronesian, Indo-European, and Dravidian. As such SLM can shed light on issues of potential universality and language specificity in contact-induced language change (Ansaldò [2005](#), [2008](#), [2009a](#)). The relationship between the three language groups involved in its evolution was similarly of an altogether different type than the better-known scenarios of exploitation, slavery, intermarriage, etc., and thus provides us with contrastive material relevant to understanding typological variation as a factor in contact-induced change. From the linguistic point of view, as I will show, SLM is an ideal case study for illustrating the need to go beyond the currently accepted frameworks of CLF.

Contrary to recent analyses, which have taken into account both socio-historical information and structural data, and which view SLM as a mixed language of trilingual base, with lexical items predominantly from contact Malay varieties and grammatical features from Sinhala and Tamil (Ansaldò

2005, 2008, 2009a; Ansaldo and Nordhoff 2009), previous accounts have traditionally considered SLM a Creole (Smith, Paauw and Hussainmiya 2004) and/or an ‘intertwined’ language (Bakker 2000). It is referred to as a Creole in an early account by a historian (Hussainmiya 1987) and is listed as Sri Lankan Creole Malay in the *Ethnologue* (Gordon 2005). Besides suggesting an inaccurate evolutionary scenario (Ansaldo 2008), classifying a variety as a Creole in fact has a significant impact on the type of shift that may occur in a community. The awareness of speaking a ‘corrupt’ or ‘broken’ variety, as is often implied in the popular understanding of the term Creole, may lead to a perception of a linguistic variety as not being ‘good’ enough, and may further strengthen a community’s desire to shift to a more standard variety (Garrett 2006: 180f.; Lim and Ansaldo 2006a, 2007). In addition, claims of interethnic admixtures are rejected by the SLM community, which has a long history of predominant endogamous practice, as can be confirmed by archival and historical sources (Ansaldo 2008; section 6.3 of this volume).

In this chapter, we will see how SLM is a completely natural and logical outcome of the ecology in which it evolves. I argue that SLM is the product of a trilingual ecology in which a significant restructuring occurs due to typological pressure, with a certain element of retention of the ancestral language due to its ethnolinguistic vitality. In illustrating the interplay between internal and external ecological factors in the evolution of SLM, I first sketch the general ecology within which SLM has evolved. In doing this, I highlight the socio-political context of the Malay Diaspora in relation to the dominant, native ethnic communities of the island, namely the Sinhalese and the Tamils. I then move on to the internal make-up of the typological matrix (TM, see section 5.2.3.3) and look at patterns of frequency that explain the grammatical output of this contact situation. As we will see, congruence, or type frequency, is the dominant factor behind the processes of selection and replication observed in the evolution of SLM. In section 6.3, I contrast this approach to previous analyses of SLM; this is useful as it reveals how exceptionalist approaches to CLF (chapter 4) can actually mislead our understanding of the evolution of new varieties. In the final section, I preview the concept of identity alignment, which will be fully addressed in the next chapter.

## 6.1 The SLM community

The evolution of the community of Sri Lanka Malays arose through one of the central practices of Western colonization, namely, the relocation of subjects from one colony to another. The Dutch and subsequently the British, in their East Indies dealings, moved large numbers of peoples from their origins in the Southeast Asian Malay world to other locations in their possession, in the case of the Malays, to Sri Lanka (then Ceylon). This way, sizeable communities of

Malays were settled on the island through various waves of deportation (as discussed in [chapter 2](#)). While the community based in the district of Slave Island may well have been there during Portuguese rule (until 1656), the bulk of the Malays trace their ancestry to exiled nobility and political dissenters, slaves, and soldiers during Dutch rule (1656–1796), and the continued importation of slaves and soldiers for the garrison during British rule (1796–1948) (Hussainmiya 1987). From the late 1800s, census data show that the Malays consistently comprised approximately 0.33 per cent of the population. The largest group of people attributed a Malay origin came as soldiers from disparate places such as Bali, Java, Riau, Ambon, the Moluku, and peninsular Malaysia (see map 3.1), such that “almost all the major ethnic groups from the region of the Eastern archipelago were represented” (Hussainmiya 1987: 48). Unlike the Hokkien traders or Portuguese adventurers ([chapters 2, 3 and 7](#)), these soldiers did not come alone; historical accounts compiled under Dutch and British rule are clear regarding the common practice of ‘Malay’ womenfolk coming alongside the soldiers in the early days of the community (see Schweitzer *ca.* 1630 in Raven-Hart 1953; van Goens *ca.* 1670 in Reimers 1932; Sourjah 2005). Overall, it has been suggested that three different communities could be distinguished: (a) a rather sophisticated diaspora of noblemen and their families (Hussainmiya 1987); (b) a ‘Malay’ garrison, what would become the Ceylon Rifle Regiment under the British (cf. Ricklefs 1974); and (c) a group of convicts of whom little seems to be known. As far as the first two groups are concerned, contacts between the groups were indeed quite frequent, due, among other reasons, to the practice of employing noblemen as officers of the troops as well as intermarriages (Hussainmiya 1987, 1990; Ansaldo 2008).

It is important to bear in mind that, under the Dutch, SLM communities were established in different parts of the island. Besides Colombo, the capital, communities of Malay/Indonesian people arose in Galle (Southwest), Jaffna (North), and Trincomalee (West) from groups placed there as soldiers, plantation labourers, and servants. A group of runaways and dissidents is also known to have established itself in Kandy at the side of the Kandyan king (see Bichsel-Stettler 1989). These communities continued to flourish under the British and are to a large extent still to be found today. Traditional and oral history sketches the SLM communities as distinct yet integrated with the other ethnic groups, namely the Sinhalese and the Tamils (Hussainmiya 1990; Saldin 2003). It must be noted that, throughout Dutch and British rule, all these groups were on roughly equal terms as colonial subjects. While the dominant Sinhalese certainly counted on prestige derived from being the original majority group of the island, parts of the Malay/Indonesian communities appear to have been highly regarded by the colonial power, based on their employment as military and police officers in particular during British rule. While contacts between SLM people and Tamils were not as frequent as previously thought (see [section 6.3](#)), the Tamil Moors



had a religious affinity with the SLM people in that both groups shared the Muslim faith. Nowhere in the history of the island do we find indications of ethnic tensions involving the SLMs and we must therefore conclude that the SLM communities were well integrated within Sinhalese and Tamil communities. As testament to the integration in their new country, when the British took over the island and made arrangements for the repatriation of all the Malay soldiers in 1799, the Malays as one refused to be repatriated because they had settled with their families and made Ceylon their home (Burah 2006: 72–3).

### 6.1.1 *Defining the matrix*

The early Malay communities in Sri Lanka were characterized by multilingualism, involving languages such as Javanese, Ambonese, Riau Indonesian, Malay, etc. (Ansaldò 2008; and section 6.3 below). As already discussed in chapters 2 and 3, in such environments a lingua franca such as Bazaar Malay was commonly used as a language of interethnic communication (see also Hussainmiya 1990: 47).<sup>3</sup> A majority of the Malay deportees actually either came directly from the Dutch fort of Batavia (established in 1619) or at least resided in and around Batavia for a prolonged period of time, before being moved on to Sri Lanka (de Haan 1922). According to de Haan, this may explain why the Malay community under the Dutch was – albeit socially and geographically of heterogeneous origins – nonetheless a close-knit, united group, with a common lingua franca. We can therefore assume that a Pidgin-derived Malay (PDM, section 3.2) variety of seventeenth-century Batavia was brought over to Sri Lanka as the language of the Malay Diaspora.<sup>4</sup>

In the ecology of Sri Lanka, this Malay lingua franca came into contact with two adstrates: colloquial Sinhala and Lankan Tamil. The former was, and is to this day, the dominant language of the majority of the population, the Sinhalese, who see themselves as the legitimate inhabitants of the island. The latter was the language of the largest minority group spoken by, amongst others, plantation workers in Southern India and Sri Lanka. Also, Tamil was a widespread language of trade throughout Monsoon Asia (see chapter 3). Dutch was never particularly influential: Dutch authorities always resorted to a variety of colloquial or trade Malay in their colonial dealings (Taylor 1983), a practice which would be later adopted by the British as well. The analysis of the typological matrix will therefore focus in particular on the interaction between prototypical features of the PDM type (see section 3.2), Sinhala and Tamil.

### 6.1.2 *SLM communities at present*

The Sri Lanka Malay Diaspora is at present distributed in various communities located around the island (see map 6.1), which vary in socio-economic



Map 6.1 Malay communities in Sri Lanka

and educational status, and their linguistic repertoire and communicative practices, as summarized in [table 6.1](#) (adapted from Ansaldo 2008). Although SLM appears to comprise perhaps some five different varieties (Lim and Ansaldo 2006a), and there is also a sense of separate identity for each of the different communities (SLM community p.c. 2003–2006), the speakers nonetheless identify themselves as the same ethnic group. The Malay Diaspora

Table 6.1. *Sri Lanka Malay communities – sociolinguistic profile*

| Community                    | Characteristics                                                                                                                                                                                                                                       |
|------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Colombo                   | Middle-upper class community in capital city; restricted usage of SLM in old-middle generations; common Sinhala (and some Tamil) competence; English fairly fluent to native speaker competence; standardizing in Malay; no SLM in younger generation |
| 2. Slave Island              | Lower class community in a poor district of Colombo; strong Tamil influences; no English                                                                                                                                                              |
| 3. Kandy and other upcountry | Middle-lower/rural class communities in the central hill country area; SLM in old-middle generations, and in some younger generation; Sinhala competence; some English proficiency, especially in younger generation                                  |
| 4. Hambantota                | Community on the south coast, traditionally heavy Sinhalese-speaking area; SLM in old-middle generations; often trilingual with Sinhala and Tamil; limited English                                                                                    |
| 5. Kirinda                   | Fishing community on southeast coast; SLM dominant in all generations; fully trilingual with Sinhala and Tamil, especially in middle-younger generations; English limited to a few individuals                                                        |

was already characterized as a close-knit community during colonial rule; the contacts between the different Indonesian ethnicities as well as the different social extractions were maintained in particular through the ranks of the army as well as through common religious practice (Ricklefs 1974). The various communities do, however, show different degrees and types of multilingualism, depending on their locality and network, as will be seen in the following sections.

One aspect that should certainly be noted is how each SLM community is characterized by some degree of multilingualism; we should therefore not conceive of SLM speakers – past or present – as uniquely associated with SLM. As argued in Lim and Ansaldo (2006a, 2007), SLM communities then and now are not to be defined by the SLM language alone but also by being multilingual. SLM varieties have until recently been native languages of diasporic communities, in the first instance independently of the knowledge of any of the source languages. As an example, children in Kirinda grow up speaking first of all Java, the SLM variety of Kirinda (Ansaldo 2005, 2008), and only later get exposed to Sinhala, through social interaction, and to Tamil in

school.<sup>5</sup> The population is trilingual to varying degrees of competence or at least semi-bilingual in SLM and Sinhala or Tamil. Even if some SLM communities are geographically quite remote, there is no linguistic isolation from the adstrates. As this chapter will show, grammatical admixture under typological congruence is high. SLM varieties could be seen as conventionalized mixed languages, i.e. varieties that emerge in situations of multilingualism alongside one of their source language(s), such as Ma'á and Media Lengua (cf. Matras and Bakker 2003: 3). However, unlike these, SLM varieties are not typical of predominantly mixed households. In more isolated communities such as Kirinda, in fact, though intermarriage exists, it is very limited and reflects a recent trend (Ansaldò 2008). We can therefore say that the SLM communities have, throughout their history, undergone a process of identity alignment in which identities are always multiple (or simultaneous) and constantly changing (Woolard 1999: 20–1; Djité 2006). This, as I will argue in [chapter 7](#), is crucial for our understanding of the why of CLF (see [section 7.2](#); Ansaldò and Lim 2006; Lim and Ansaldò 2007).

## 6.2      Selection and replication in SLM<sup>6</sup>

From the general ecology sketched in [section 6.1](#), we derive the principal features that constitute the typological matrix: (i) PDM features as the original vernacular language; (ii) Colloquial Sinhala features as the dominant, prestige language; and (iii) Colloquial (Lankan) Tamil features as an influential ethnic group.<sup>7</sup>

In what follows, I examine the nominal and verbal systems of SLM, focusing especially on two salient aspects: (1) the case system, and (2) the tense and aspect (TA) system. The variety of SLM described here is Kirinda Java (see Ansaldò 2005, 2008). It is an ideal variety in that it is associated with a prototypical SLM community which, unlike most others, has remained immune from the most recent socio-political changes that much of Sri Lanka is undergoing. These changes, which have occurred during the last three decades of independence, are reflecting changes in the linguistic ecology of SLM communities, as can be seen in the shift and endangerment they are undergoing (Lim and Ansaldò 2006a, 2007; Ansaldò 2008). The village of Kirinda (see [map 6.1](#)) has been relatively sheltered due to its geographical remoteness; at the same time, it is situated in a typical SLM ecology, with cultural and economic ties to Tamil groups (religion and fishing) and general contacts with Sinhalese neighbouring groups. Also, it is the only fully vital community in which children grow up speaking SLM as their mother tongue.<sup>8</sup>

Table 6.2. *Etymology of case markers in SLM*

| Case                  | Marker        | Etymology                                     |
|-----------------------|---------------|-----------------------------------------------|
| Dative                | = <i>nang</i> | Malay <i>nang</i> 'towards'                   |
| Accusative            | = <i>yang</i> | Malay <i>-nya</i> [+def marker]               |
| Possessive            | = <i>pe</i>   | PDM <i>pe</i> [from Possessive <i>punya</i> ] |
| Locative              | = <i>ka</i>   | Malay directional <i>ke</i>                   |
| Instrumental/Ablative | = <i>ring</i> | Jakarta <i>bikin</i> 'to make'                |
| Comitative            | = <i>le</i>   | (Malay <i>de &lt; dengan</i> 'with')          |
| Nominative            | Ø             |                                               |

### 6.2.1 Case

SLM shows a full set of post-nominal case markers, a unique feature among Malay restructured varieties, considering that Malay languages are typically isolating to mildly agglutinative and do not feature nominal case morphology. On the other hand, case systems are common in Indo-Aryan grammars, to which Sinhala belongs, as well as in Dravidian languages, to which Sri Lankan Tamil belongs. Both languages show agglutinative morphology with fusional tendencies. Moreover, because of over a millennium of intense contact, Sinhala and Tamil grammars have converged typologically and show substantial similarities (Masica 1976; Emeneau 1980). As the rest of this section shows, SLM has developed agglutinative morphology with incipient fusional tendencies in the nominal system, suggesting a typological shift away from the isolating type, as briefly illustrated in (1) (and see Ansaldo and Nordhoff 2009):

- (1) *Ni aanak=nang      baik      buku=yang      attu-aada*  
 This student=DAT    good    book=ACC    one=have  
 'This student has a good book'

As we can see in the example, Dative and Accusative are marked through suffixes attached to the noun. Note also the verb-final word order, which follows Sinhala and Tamil (or 'Lankan') typology, and not Malay/Indonesian, which is strongly verb-medial.

#### 6.2.1.1 Morpheme sources

As is typical of contact languages of the Malay world or PDM varieties (Adelaar and Prentice 1996), the lexicon of SLM is heavily influenced by material of generic Malay origin. Table 6.2 lists the SLM case markers and their probable etymology:<sup>9</sup> as can be seen, the lexical origin of these grammatical markers is clearly Malay.

Table 6.3. *Functions of case in adstrates Sinhala and Tamil*

| Sinhala               |                         | Tamil                    |                                        |
|-----------------------|-------------------------|--------------------------|----------------------------------------|
| Case                  | Function                | Case                     | Function                               |
| Nominative            | Agent (+Animate)        | Nominative               | Agent                                  |
| Dative                | Experiencer             | Dative                   | Experiencer                            |
|                       | Goal, Beneficiary       |                          | Goal, Beneficiary                      |
|                       | Possessor               |                          | Possessor                              |
| Accusative (optional) | Patient (Human/animate) | Accusative               | Goal                                   |
| Genitive-Locative     | Location                | Genitive-Locative        | Location Path/<br>Temporary possession |
|                       | Temporary possession    |                          |                                        |
| Instrumental-Ablative | Instrument/Source       | Ablative                 | Source                                 |
| Vocative (+ANIM)      | Addressee               |                          |                                        |
| Comitative            | Association             | Comitative               | Association                            |
|                       |                         | Associative/Instrumental | Association/Instrument                 |

Table 6.4. *Thematic roles in SLM and adstrates Sinhala and Tamil*

| Case         | SLM               | Sinhala       | Tamil          |
|--------------|-------------------|---------------|----------------|
| Nominative   | Agent             | ✓             | ✓              |
| Dative       | Goal              | ✓             | ✓              |
| Accusative   | Patient           | ✓ if +Animate | ✓ if +Definite |
| Possessive   | Possessive        | ✓             | ✓ + Location   |
| Comitative   | Associative       | ✓             | ✓              |
| Instrumental | Instrument Source | ✓             | Ø              |

### 6.2.1.2 Structural features of case in SLM, Sinhala, and Tamil

As this section shows, syntactic and semantic features of SLM are predominantly of the Lankan type. Table 6.3 shows the functions of case in the adstrates Sinhala and Tamil, while table 6.4 compares the functions of SLM case with them.

Comparing tables 6.3 and 6.4, we can make the following observations:

- i. Prototypical Agents in SLM are unmarked, as in the adstrates Sinhala and Tamil.
- ii. Experiencers and Goals in SLM are marked identically; they correspond to Dative case in the adstrates. Lack of Volition or Control, just as in Sinhala, is the key semantic feature here.
- iii. Accusative marking parallels the adstrate typology. The SLM definite object marker (ACC) shares the feature [+definiteness] with Tamil and is optional as in Sinhala.
- iv. SLM shares the Instrumental-Ablative syncretism with Sinhala.
- v. In SLM, Genitive and Locative receive different marking.
- vi. Comitative shows complete congruence.



Figure 6.1 The Control feature

Table 6.5. *Assignment of Dative in SLM and adstrates Sinhala and Tamil*

| T-role              | SLM | Sinhala | Tamil |
|---------------------|-----|---------|-------|
| Experiencer         | ✓   | ✓       | ✓     |
| Goal/Benefactive    | ✓   | ✓       | ✓     |
| Possession/Location | ✓   | ✓       | ✓     |

It is clear that the SLM language is a hybrid system in which Malay lexical items combine with Sinhala and Tamil grammatical categories. How this system evolves is discussed below.

6.2.1.3 *Functional alignments*

While syntactically case is realized post-nominally in SLM, which is congruent with Lankan grammar, semantically there are aspects in which it can be considered innovative. In what follows, I focus on the core cases of SLM and the grammatical relations they mark. Arguably the most prominent function of case in Sinhala and Tamil is the marking of Control, as can be seen in the high frequency of ‘Dative subjects’ or ‘Experiencers’ (see [figure 6.1](#)).

This alignment is clearly present in the grammar of SLM: Dative case is assigned as in the adstrates Sinhala and Tamil, as shown in [table 6.5](#).

[Table 6.5](#) illustrates a typical case of South Asian Dative, in which several functions conflate; this is defined by the distinction between Agents, who are in Control or volitional, and Actors, non-volitional Agents who do not control the action, as is the case for psych verbs, perception verbs, and modality (examples 2 and 3).<sup>10</sup>

- (2)

*Go=dang*

I=DAT

‘I see Mr Jalaldeem’

*Mr. Jalaldeem=yang*

Mr. J=ACC

*kutumun*

see
- (3)

*Sir*

teacher

‘The teacher sent the children home’

*aanak=pada=yang*

child=PL=ACC

*ruma=nang*

house=DAT

*e-luppa*

PST-send

Table 6.6. *Assignment of Accusative in SLM and adstrates Sinhala and Tamil*

| SLM          | Sinhala      | Tamil                     |
|--------------|--------------|---------------------------|
| + DEF        | +/- DEF      | + DEF                     |
| +/- ANIM     | + ANIM       | +/- ANIM                  |
| – obligatory | – obligatory | + obligatory on<br>+ ANIM |

The typological congruence between Sinhala and Tamil Dative is so strong that SLM replicates these functions identically. Dative is a very robust feature in SLM; for most speakers it is obligatory, particularly when in the function of Experiencer as opposed to Goal (or Beneficiary). The Agentive–Experiencer/Goal opposition – i.e. zero-marking which appears to be reserved for prototypical Agents, and the Dative marker which covers first arguments in non-agentive roles as well as a range of Patient/Goal roles – is the most consistent marking found in the noun phrase (NP), based on frequency and distribution.

Turning to the functions of Accusative in SLM, we find a less congruent picture, as shown in [table 6.6](#).

Definiteness is undoubtedly a function of Accusative case in both Sinhala and Tamil, though they differ in the constraints that apply. In Sinhala, Accusative is only used with animate objects, so that definiteness can only concur with Animacy. Moreover, even with animate nouns, Accusative is highly optional in Sinhala (Gair and Paolillo 1997).<sup>11</sup> We can therefore say that it is not obligatory to mark definiteness in Sinhala. In Tamil, Accusative case is only obligatory for a human direct object. With non-human referents, this case is only used to mark definite objects. In SLM, Animacy appears not to be widely marked, though the pronominal system shows variable case marking for Dative first and second person singular (4), a manifestation of Animacy distinctions captured in Silverstein’s (1976) hierarchy as shown below (see also Ansaldo and Lim 2004b; Ansaldo 2005, 2008, 2009a):<sup>12</sup>

- (4) *go=dang*      I=DAT  
*lu=dang*      you=DAT  
*dia=nang*      he=DAT  
*kitang=nang*      we=DAT  
*lorang=nang*      you.PL=DAT  
*derang=nang*      they=DAT

As in Tamil, Accusative sometimes marks definiteness in SLM; however, as in Sinhala, it is highly optional. Moreover, Accusative marking sporadically also occurs for what appears to be emphasis (5), therefore adding a novel function to its repertoire. Interestingly, in Colloquial Malay varieties, an emphatic



marker =*nya* is found that closely resembles the SLM Accusative marker =*yang* (or =*ya*). Different features are therefore combined in what can be seen as an optional marker of prevalently definite objects (Ansaldò 2005, 2008, 2009a).

- (5) *Inni kendera=yang ba pi*  
 this chair=ACC take go  
 'Take **this** chair away'

Thus, unlike with the Dative case, the partial congruence between Sinhala and Tamil Accusative case leads to a partial replication with innovative features in SLM.

Moving our analysis outside the core cases, two more interesting observations can be made. In SLM, we find a case of Instrument/Ablative syncretism, meaning that the Instrumental case marker =*ring* may indicate Source:

- (6) *Market=ring ais tra baaru ikkang billi bawa*  
 market=ABL ice NEG new fish buy bring  
 'Get me some fresh fish from the market'

This parallels the functions of Instrument in Sinhala, but not in Tamil, in which Instrument and Ablative are clearly distinct (Silva 2003; Ansaldò 2005, 2008, 2009a). A plausible explanation for this pattern could be found in the fact that, of the two adstrates, Sinhala has more speakers and more prestige; where the adstrates diverge, SLM seems to follow its patterns.

However, this is not always the case: in the realization of Possessive, a different outcome is found. While in Sinhala and Tamil, as table 6.3 shows, Location and Possession are conflated into one case (Schiffman 1999; Silva 2003), SLM has two distinct cases for Locative and Possessive. The SLM Possessive case suffix is =*pe*, a derived form of the Malay *punya* 'to possess':

- (7) *Go=ppe tumman go=yang e-tolak*  
 I=POSS friend I=DOM PST-push  
 'My friend pushed me'

Reduced variants of *punya* are well attested as distinguishing features of contact-Malay varieties (PDM) such as Bazaar Malay, Baba Malay, etc. As we will see in chapter 7, this Possessive marker is one of the most robust features of Bazaar Malay varieties, which leaves little doubt to the fact that its retention in SLM is justified by the founder effect (Mufwene 1996b). As an archaic feature of a Malay lingua franca, it is most likely that this feature was maintained from the original vernaculars of the SLM community and its exaptation led to the development of a new case that distinguishes SLM from its adstrates.

#### 6.2.1.4 Evaluation

The case system of SLM indicates that overall frequency determines selection in the typological matrix. The general congruence of Sinhala and Tamil leads

to high type-frequency of the Lankan grammatical features that are selected, as seen in morphology, constituent order and case marking.

(A) In the case of Dative, Tamil and Sinhala show basically total congruence in the typological matrix of what is a typical South Asian Dative; moreover, Dative is very high in discourse frequency, being obligatory and employed in a wide range of functions. The result can be captured as follows:

Features with high token- AND type-frequency are likely targets for identical replication in the new grammar.

(B) In the case of Accusative, however, there is less typological congruence between Sinhala and Tamil, since the former is more sensitive to Animacy than Definiteness. Therefore type-frequency is undermined. Also, Accusative appears less prominent in terms of discourse semantics, being more limited in its functional range and mostly optional. This results in low discourse frequency, as is also revealed by its absence in other analysis of SLM case (Smith, Paauw and Hussainmiya 2004). In this case, the SLM Accusative emerges as an adaptive innovation, which only partially combines dominant features of the typological matrix and also exhibits a novel function. We can therefore generalize as follows:

Features with low token- OR type-frequency are likely targets for innovative replication (or for lack of selection) in the new grammar.

Finally, in the non-core cases, we see two interesting patterns emerging. First, we note that, where there is a significant lack of typological congruence between the adstrates, as for the Instrumental/Ablative case, SLM patterns with the dominant language, Sinhala. This can be attributed to token-frequency, confirming the general pattern of frequent features being more robust in the evolution of grammar (see 5.2.2), in this case justified by prestige and numerical majority. However, this is not always the case: a different pattern can be observed in the Possessive. Here an early Possessive morpheme retained from the ancestral language – one of the most robust PDM features across contact varieties of Malay – leads to the formation of a new case marking, which I have attributed to the founder principle. These different outcomes show that typological analysis based on frequency can explain much of what goes on in the restructuring process, but not necessarily everything.

### 6.2.2 *Tense and aspect*

Unlike in the case system, tense and aspect (TA) marking in SLM does not show the same overall congruence with adstrate typology. This is not surprising in a sense, because while the systemic organization of case is tightly structured as a declension system around interdependencies of Thematic-role (T-role)

Table 6.7. *Morpheme sources for TA markers in SLM*

| Function                     | Morpheme                                     | Etymology                                                   |
|------------------------------|----------------------------------------------|-------------------------------------------------------------|
| Perfective/Past              | <i>si-</i>                                   | PDM <i>su(dah)</i> , (perfective)                           |
| Past                         | <i>e-</i>                                    | <i>si</i> > <i>se</i> > <i>e</i> ?                          |
| Perfect                      | <i>ada</i> (post-verbal)                     | PDM <i>ada</i> ‘to have’ (emphatic)                         |
| Progressive/habitual/present | <i>ar(e)-</i>                                | PDM <i>ada</i> (progressive)                                |
| Future/irrealis              | <i>ati-/ai-</i>                              | M. <i>nanti</i> (‘soon’)                                    |
| Completive                   | <i>abis</i> (post-verbal)                    | M. <i>habis</i> ‘finish’                                    |
| Continuous                   | <i>ambe.(l)dudu.(wattu)</i><br>(post-verbal) | <i>ambel</i> + <i>duduk</i> ‘take sit’                      |
| Negation                     | <i>tada</i>                                  | PDM <i>tar/tra</i> , no(t) (M. <i>tidak ada</i> ‘not have’) |

assignments (see e.g. Haspelmath 2008), tense and aspect are really different semantic domains. As aptly observed especially by Matras (2000, 2003), in contact-induced change, a whole system can be transferred from one language to the other, in particular when the system has high semantic/pragmatic visibility. This type of ‘categorical fusion’ has been documented in other contact environments, such as discourse particles in Singlish (see Ansaldo and Lim 2004a; Lim 2007; Lim and Ansaldo 2006b). The contrast between the evolution of case and that of TA in SLM also shows that tightly organized systems are likely to be selected *in toto* in the restructuring process.

While Malay varieties do not mark semantic roles through case, they do exhibit TA aspectual distinctions; the result is a more complex competition between Lankan and Malay features than observed in the case system. In both Sinhala and Tamil, tense is a more salient semantic domain than aspect. On the other hand, Malay varieties in general do not mark tense, but have elaborate aspectual systems. For these reasons, I focus on temporal distinctions in the adstrates rather than aspectual ones, as the latter domain is optional and variable. Tamil aspect, for example, is highly optional and is in most cases marked through aspectual verbs, rather than grammatical markers, and therefore not fully grammaticalized (Schiffman 1999). When we look into how tense is realized in the adstrates, we see that while Sinhala has two tense categories, past and non-past, Tamil has three, namely past, present, and future. As shown in table 6.7, SLM has over time acquired tense distinctions which follow the Sinhala pattern;<sup>13</sup> also clear in table 6.7 is the fact that aspectual distinctions still play a significant role in the marking of events in SLM.

Table 6.7 illustrates the predominantly Malayic origin of aspectual (and tense) markers in SLM. As can be seen, all TA markers can be etymologically identified in varieties of Malay, particularly PDM (see Adelaar 1991 for a detailed discussion): Ambonese Malay for the perfective and past markers,

Table 6.8. *Aspect in PDM varieties*

| Function      | Baba Malay             | Bazaar Malay      | Cocos Malay    |
|---------------|------------------------|-------------------|----------------|
| Perfective    | <i>suda</i>            | <i>suda/habis</i> | <i>(su)dah</i> |
| Incompletive  | <i>belum</i>           | <i>belum</i>      | <i>b(e)lum</i> |
| Recentive     |                        | <i>baru</i>       | <i>baru</i>    |
| Future        | <i>mo/nanti</i>        | <i>nanti</i>      | <i>nanti</i>   |
| Irrealis      | <i>mo/nanti</i>        | <i>mahu</i>       | <i>mau/mu</i>  |
| Progressive   | <i>lagi (... lagi)</i> | <i>lagi/ ada</i>  | <i>lagi</i>    |
| Past/habitual | <i>dulu</i>            | <i>dulu</i>       |                |

Moluccan, Baba, and Bazaar Malay for the present/progressive, and Bazaar and Moluccan Malay for the negation marker. The only non-evident marker is the Continuous construction *ambe(l)dudu*, illustrated in (8):

- (8) *Go market=nang (e)-pi ambe-ar-dudu si-jato*  
 I market=towards (PST)-go take-PROG-sit PFV-fall  
 ‘While I was going to the market I fell.’

The form *ambil* can be found in Indonesian Malay as a pre-verbal particle with identical function (Slomanson 2006).<sup>14</sup> As shown in table 6.8, the resemblance between SLM and other contact Malay varieties is quite strong in terms of aspectual systems and indicates a fairly clear target of selection.

While SLM has predominantly aspectual marking, it does mark at least past vs. non-past, a fact that implies a functional alignment with Sinhala. At the same time, tense in general is not obligatory in SLM and aspectual marking is clearly more frequent in discourse. Regarding aspect, considering the relatively scanty information we have on this domain in Sinhala and especially (Lankan) Tamil, it is difficult to generalize. But the lexical and structural isomorphy with PDM categories leaves little doubt that it is Malayic aspectual categories that are selected in the TM. Still, the fact that morphologically TA is more agglutinative than isolating reveals the depth of congruence with the Lankan type in the overall grammar of SLM. Note also that TA morphology in SLM is pre-verbal, while it is post-verbal in Lankan typology; this yields a certain variation in the constituent order of the SLM verb phrase (VP) which, though predominantly V-final, also allows for V-medial structures (Slomanson 2006).<sup>15</sup>

We can thus observe the following in the evolution of the verbal domain in SLM:

- (a) Sinhala and Tamil do not show overall congruence in aspect; moreover, there are PDM aspectual features in the matrix. Therefore SLM selects Malay aspectual features due to (i) low type- and token-frequency of Lankan aspectual features; this is the result of lack of typological congruence

between the adstrates, and optionality of aspect marking; and (ii) vitality of the ancestral features, which can be seen as a founder effect.

- (b) Sinhala and Tamil tense categories have a relative degree of functional congruence and are salient semantic categories in discourse. The absence of tense in PDM results in tense development in SLM along Lankan lines.
- (c) TA is a less tightly organized system than case. In addition, in the matrix in which SLM evolves, there is competition between vernacular and adstrate features in the verbal domain not found in the nominal domain. These two aspects combined yield the typological asymmetry that characterizes SLM grammar.

The well-developed aspectual system of PDM varieties obviously offers plenty of material for selection in the matrix from which SLM derives its features, and it is indeed aspectual categorization that dominates SLM grammar rather than tense distinctions.

In terms of constituent order, the VP shows less congruence with Lankan grammar than the NP.<sup>16</sup> The only way to explain the contrast between the N and the V domain lies in the notion of typological ‘gang-up’ (Ansaldò 2005, 2008, 2009a): it is the congruence of the two adstrates, combined with the lack of PDM marking of grammatical relations, that leads to an overall restructuring of the NP towards Lankan type in SLM. Conversely, it is the lack of adstratal congruence, together with the presence of PDM aspectual material, that leads to limited restructuring of the VP in the SLM TM. This shows that any plausible explanation of the evolution of SLM must be derived from the observation and analysis of the trilingual situation, in which all the three languages present in the TM are given serious consideration. As we will see below, this corroborates the fact that previous suggestions that rely on bilingualism alone (Slomanson 2006: 156) and the overwhelming influence of Tamil (Smith and Paauw 2006) need to be revised (see also Ansaldò 2008, 2009a; Ansaldò and Nordhoff 2009; Nordhoff 2009a).

### 6.3 Freeing SLM from the chains of exceptionalism

Even if the study of SLM over the past two decades has generated only a small body of literature addressing limited aspects of the relevant speech communities, *a priori* assumptions about the nature of language contact as well as expectations derived from specific ideological stances have led to the construction of a paradigm that aims at perpetuating itself, without further questioning its own evolution (cf. Ansaldò 2005, 2008). This offers us a fascinating opportunity for reflection on the role of ideology in theory construction as well as the limiting role that such theories have in analysis. It also shows us the necessity of novel approaches to the study of contact language formation.

Being in the fortunate position of collecting first-hand data over a prolonged period of time and documenting oral history from members of various SLM communities across the island, I address the issues that have mistakenly become accepted wisdom and discuss the reasons for the current state of affairs.

As mentioned at the outset of this chapter, SLM has been described as a Creole that developed as a result of intermarriage between Malay men and (Tamil) Moor<sup>17</sup> women (Smith, Paauw and Hussainmiya 2004) or close contact between SLM speakers and Tamil Moors (Smith and Paauw 2006). This view is based primarily on the historical observations of Hussainmiya (1987, 1990) regarding the records of marriage under the Dutch (*Thombas*) which, he suggests, show several cases of intermarriage between SLM and Tamil Moors; the religious affinity between the Moors and Malays, both sharing the Muslim faith, is also taken to support such alliances.

One might want to be suspicious of the obvious incongruence between the idea, on one side, of bilingual admixture and, on the other side, ‘creolization’, as the two are not considered the same in language contact studies. More importantly, the notion of creolization as a special diachronic development has been extensively criticized and proven wrong in a number of recent works (Mufwene 2000, 2001; Ansaldo, Matthews and Lim 2007). It is in fact the idea of bilingualism that emerges as the most entrenched, *a priori* assumption in the evolution of SLM, which leads to analyses that often disregard the role of Sinhala or at least favour closer inspection of Tamil grammar in the search for ‘substrate’ evidence (Bakker 2000; Slomanson 2006; Smith and Paauw 2006). Bakker (2000) in particular suggests that SLM should be seen as a mixed language that evolved ‘rapidly’ from a hypothesized earlier reduced linguistic code. As has become clear in the previous sections, however, neither historical nor typological evidence lends itself to a bilingual scenario of genesis, and ignoring the multilingual matrix is methodologically unsound in an evolutionary scenario (chapter 5). How then did such a bias become justified? I have referred to this dominant SLM paradigm as the ‘Tamil bias’ (Ansaldo 2008), summarized in the following section.<sup>18</sup>

### 6.3.1 *The Tamil bias: construction and deconstruction*

In what follows, I retrace and critically evaluate the claims that have led to the construction of the Tamil bias in the evolution of SLM (Ansaldo 2005, 2008):

- i. Hussainmiya (1987), the first, extensive account of SLM history, notes that the Dutch *Thombas* (records of marriage kept under the Dutch) list some mixed marriages, including Malay–Tamil Moor, Ambonese–Malabarese, Malay–Sinhalese; he also suggests that SLM may be influenced by Sinhala, Tamil, or both.<sup>19</sup>

- ii. Hussainmiya (1990) notes frequent episodes of religious and cultural exchange between Malays and Moors.
- iii. Based on the above, Bakker (2000) argues for rapid typological convergence and admixture with Sinhala and Tamil as a recent development; he assumes that 'creolization' may account for earlier stages of SLM's history and postulates the evolutionary path from a Malay–Tamil jargon to contemporary SLM. Bakker (2000) rests on the assumption of parallel evolution of Sri Lankan Creole Portuguese (SLCP) and SLM.
- iv. Based on (ii) and (iii), Smith (2003) engages in an investigation of Accusative markers in SLM *in order to prove the substantial influence of Tamil*. The paper concludes that while positive Tamil influence cannot be proven, lack of Sinhala evidence could be argued for.
- v. Smith, Paauw and Hussainmiya (2004) present a slightly revised position claiming both influence from Tamil and general Lankan influence in the evolution of SLM.
- vi. Based on (iii), Heine and Kuteva (2005) simply present SLM as a Tamil–Malay Creole as if it were a given.
- vii. Based on (iii), Winford (2005) suggests that SLCP might be the result of Tamil speakers acquiring Portuguese.
- viii. Smith and Paauw (2006) argue for a Tamil-like tense, mood, and aspect (TMA) system in SLM.
- ix. Slomanson (2006) argues for convergence between SLM, Sinhala, and Tamil; Malay–Tamil bilingualism is still invoked to explain aspects of tense marking.

In order to assess the original historical claims found in Hussainmiya's work, I investigated the same records of marriage referred to in the historical sources. The results, discussed in Ansaldo (2008), show that the records for the period up to 1796 are damaged by water, making significant parts of the entries illegible. The most revealing information for identification here are the signatures of the parties. There is however hardly any information on ethnic groups, which makes it difficult to identify Malay/Indonesians and Moors given that both groups share the practice of adopting Arabic names. In a particularly interesting section in the *Thombas* dedicated to mixed marriages (cf. Hussainmiya 1987), only 5 of 238 entries clearly refer to individuals of Javanese origin: of these, two records refer to Javanese–Moor marriage, one to a Javanese–Javanese marriage, and the remaining two are unclear. In the following period until 1919, a majority of marriages involve Western parties; there are two clear entries involving Malays, one married to a Eurasian (between 1867 and 1897), and one to a Burgher (1885–1897). From 1897 onwards, race is clearly specified; of 196 entries only one is Malay. Clearly, this evidence is not significant to support a claim of bilingual admixture nor of privileged contacts with the

Moors on the part of the Malays, a fact that is fully corroborated by the oral history collected in 50 Malay families of Sri Lanka (Ansaldò 2008).

To be fair, in (i) and (ii) there was only a weak suggestion of a possible Tamil influence in the evolution of SLM; however, subsequent linguistic literature has stubbornly developed this claim and turned it into a truism, notwithstanding the fact that direct attempts at proving it right have basically failed. The parallel with the Portuguese Creole of Batticalhoa expressed in (iii) is likewise devoid of historical accuracy. SLCP is the variety of Indo-Portuguese developed in Batticalhoa, a traditionally Tamil-inhabited territory on the Eastern coast of Sri Lanka. This variety, like SLM, displays strong influence from Sinhala/Sri Lankan Tamil typology, but may well show stronger influence from Tamil than Sinhala. However, it cannot be assumed that what may hold for the genesis of SLCP may be extended to the history of SLM, as the ecology in which the two varieties evolved is markedly different (see also [chapter 8](#)). SLCP may have developed in a small, focussed community in contact with predominantly Tamil-speaking communities, where predominantly Portuguese males would have been in contact with Tamil speakers, though Sinhala influence has not been ruled out. SLM on the other hand evolved in Malay-speaking communities, including women, in Sinhala- and Tamil-speaking areas of Sri Lanka in more diffuse settings. Therefore while SLCP may be a result of Portuguese–Tamil admixture through intermarriage, in the case of SLM, intermarriage, and admixture with specifically Tamil speakers is neither attested nor historically plausible. It is crucial to consider that, while intermarriage was a common, quasi-institutionalized practice in Portuguese domains overseas (Bethencourt 2005; [chapters 2, 3 and 7](#) of this volume), intermarriage was clearly discouraged in SLM communities, especially with Tamil Moors who had low social status. Postulating the development from an early jargon is likewise historically unfounded and theoretically problematic: why would the communities of Malays, socially fully functional as soldiers, guards, planters, and fishermen speak a ‘jargon’?

The most serious attempt to date to argue for Tamil as a primary ‘substrate’ is behind claim (iv) and, crucially, ends up failing to find the evidence. The fact that absence of Sinhala is found, rather than presence of Tamil influence, is not surprising at all considering that its focus is on finding proof of Tamil and not Sinhala influence. The same methodological problem is found in a more recent attempt that underlies the claim (viii) of a Tamil-based TMA, as only Tamil is investigated and Sinhala data are absent.<sup>20</sup>

As shown in [section 6.2](#), there are clear functional-typological reasons why the SLM Accusative (partly used to mark definiteness) looks the way it does. It is the result of low type/token-frequency and not surprising in a typological sense. Likewise I have shown that the aspectual system of SLM is mostly replicated from PDM, a conclusion also reached in a different, generative approach in Slomanson (2006). Also, the analysis of tense presented above argues for



congruence with Sinhala, not Tamil, in SLM's basic tense categories, as confirmed for the Kandy variety of SLM in Ansaldo and Nordhoff (2009). Claim (vii) suggests that SLM is the result of Tamils and Sinhalese trying to learn a Malay lingua franca but – failing to do so – producing a Lankan version of it. That they would want to do so in the first place is simply sociolinguistically untenable and there is no historical evidence to back this up. Claims (v) and (ix), though clearly more balanced, still suffer from a preferential treatment of Tamil data.

Considering the overwhelming linguistic and historical evidence in favour of trilingual admixture, it seems fair to suggest that the only way to explain the development of such a paradigm lies in theoretically biased preconceptions of CLF. Whether driven by the wish to prove 'intertwining' or 'creolization', researchers must have assumed that only an exceptional genesis could produce something like SLM. This exceptional genesis would have required one or several of the exceptional features deconstructed in chapter 4, namely pidgin-to-creole development, bilingual creation by children raised under restricted input, or the more opaque 'creolization'. Below I briefly review these ideas in relation to the evolution of SLM.

### 6.3.2 *SLM and exceptionalism*

Though there is good reason to believe that 'creolization' does not refer to a specific diachronic process nor to a specific set of extraordinary sociolinguistic conditions (see Mufwene 2001, 2008; DeGraff 2003, 2005; Ansaldo, Lim and Mufwene 2007), in what follows I will treat the possibility of applying established views of creolization and admixture to the evolution of SLM.

*I. SLM as creolization.* If we view creolization as a predominantly socio-historical concept of disenfranchised vernaculars developed in European plantation environments (e.g. Mufwene 2000, 2001), we clearly encounter problems of historical accuracy and social nature. SLM varieties are not products of such conditions, as clearly illustrated by their history (see section 6.2). Moreover, the relationship between Sinhala, Tamil, and Malay cannot be compared to a superstrate–substrate situation. Sinhala and Tamil would have been adstrates present from the start, on a par, in a sense, with colloquial and contact Malay varieties, as the languages of colonial subjects. Sinhala was, and is, the numerically dominant one, while Tamil is a strong minority language. Many SLM speakers of the middle and older generations are trilingual in SLM, Sinhala, and Tamil, and oral history tells us this has been the norm for a long time. Only a very partial interpretation of this view, namely that of disenfranchised vernacular, would allow one to refer to SLM as a Creole.

Let me briefly add that, as discussed in Ansaldo (2005, 2008), other exceptionalist scenarios do not work for SLM either. SLM grammar is difficult to

account for within current views of relexification (cf. Lefebvre 1998, 2001): its lexicon is mixed yet derived predominantly from numerous varieties of the Malay/Indonesian region; syntactic patterns show significant adstrate influence, but we also find Malay features, as for example in aspects of its verbal morphology (cf. Adelaar 1991). In this sense the mixture of SLM grammar appears to be a product of selection and replication from a multilingual matrix, as argued in section 6.2 (see also Aboh and Ansaldo 2007). More importantly, the situation in SLM is the reverse of the typical relexification scenario, in combining ‘new grammar’ with ‘old lexicon’ (Bakker 2000).<sup>21</sup> SLM does not fall under the profile of the Creole Prototype either:<sup>22</sup> it is surely not a ‘young’ grammar in the sense of McWhorter (2004), as it presents us with elaborate nominal and verbal morphology that – without further elaborating on the still somewhat subjective notion of complexity – does not appear young in McWhorter’s sense and does not look ‘simple’ when compared to its putative lexifier.

*II. SLM as rapid admixture.* Assuming that admixture applies to a specific processes of bilingual restructuring (e.g. ‘intertwining’, Bakker 2000), it is clear that bilingualism is not sufficient to account for the evolution of SLM and partial trilingualism has to be acknowledged. Note also that it is unclear whether a sharp distinction between admixture and creolization can be drawn (Grant 2007); Creole grammars such as Sranan, for example, are indeed more ‘mixed’ than previously acknowledged. Invoking a rapid development from a reduced linguistic code, as I have already shown, is ahistorical and atheoretical, based on the discussion presented in chapters 3 and 4. As shown in Ansaldo and Nordhoff (2009), SLM scores very high in terms of ‘maturity’ (Dahl 2004) or age (McWhorter 2005), as it presents us with overwhelming grammatical features that can be seen as products of accumulation over time, such as case (Dahl 2004; see also section 6.4 below). Moreover, the ‘heavy Tamil pressure’ invoked in Bakker (2000) to account for rapidity is implausible in the modern history of Sri Lanka to date.<sup>23</sup> If there were ever a linguistic type of pressure in the recent history of the island, this would have come from Sinhala, considering the political status it was afforded after independence, as seen in the Sinhala-only language policy. Moreover, considering the negative association that the Sinhalese majority has developed for Tamils (and the Tamil Tigers, regarded as a terrorist separatist group), there would be an additional reason for a minority such as the SLM not to want to be identified as Tamils, i.e. not to speak their language. Finally, as argued in chapter 3, reconstructing a pidgin ancestor is an aspect of the biased exceptionalist paradigm that demands a monogenetic, linear view of language evolution (see also chapter 8). This view is inappropriate in the multilingual ecologies in which CLF takes place (Ansaldo and Matthews 2007; Den Besten 2007) and unnecessary in the evolutionary framework adopted here. So, if

admixture had to account for the evolution of SLM, it would have to be trilingual, and gradual.

I limit myself to these conceptual frameworks as I believe they are, at this stage, the most influential ones within theories of contact language formation in particular in accounting for CLF of the type seen in SLM. One might still want to view SLM as a ‘special’ case of second-language acquisition (SLA) and thus link it to ‘creolization’ as restricted SLA. I seriously doubt this to be the case: there is no historical evidence of restricted input in the formation of SLM (and restricted input is considered implausible in a multilingual ecology, as discussed in chapter 5). Nor, as noted above, is there evidence of an early grammar stage that could have been material for reanalysis in attempting to acquire a target language. Surely the Malays did not create SLM by ‘failing’ to acquire Tamil or Sinhala, because they did indeed acquire Sinhala and Tamil quite fluently (see [section 6.2](#)).

### 6.3.3 *The normal evolution of SLM*

In many multilingual communities around the world, where multilingualism is not institutionally supported through schools, education, etc., one of the languages may find its domains of usage shrinking over time. For example, if one of the codes used becomes limited to the home domain, it typically grows weak, and interference from the more frequently used languages is expected. This essentially is language attrition and also shows the following characteristics (Thomason and Kaufman 1988):

- There is gradual abandonment of the ancestral language (AL).
- It is typical for minority groups in (a) colonization or (b) nation-expansion processes.
- It is typical of stigmatized linguistic codes.
- It happens in the transition from monolingualism in AL to multilingualism in L2/L3.
- L2/L3 take over domains of usage.
- AL speaker numbers gradually decrease and so does competence in AL.
- If AL has vitality, some features survive; if not, death occurs.

In the transition process described here, L2/L3-dominant individuals lead the change through intense code-mixing, structural, and lexical transfer. As noted in Haugen (1972: 37), the speaker “dismantles and reorders the language he already knows”. The motivation is the AL speakers’ increasing accommodation to L2/L3 under strong social and structural pressure through increased multilingual competence (Ansaldo 2005, 2007; Ansaldo and Lim 2006). In rare contexts where AL is vital, a partial maintenance is indeed possible (e.g.

Anglo-Romani, Thomason and Kaufman 1988), typically in cases where the change is gradual (Winford 2003).

This is a most plausible scenario for the genesis of SLM, which does not require exceptionalist explanations but relies on well-established sociolinguistic observations. The SLM communities were and are characterized by a solid integration within the larger communities of Sinhala and Tamil speakers, as their high degree of trilingualism to this date shows. At the same time, however, they have a very strong sense of their own SLM identity, as testified by their vital cultural, grass-root organizations as well as the active engagement in their own linguistic practices, seeking to improve the status of their language (Ansaldo and Lim 2006; Lim and Ansaldo 2006a, 2007). Moreover, such an account is in complete harmony with the evolutionary framework presented below (and the related functional-typological data analysis), where input is never ‘impoverished’, replication is never ‘imperfect’, and altered or innovative replication is expected to be the norm in multilingual, informal transmission while identical replication is specific to normative, monolingual ecologies (chapters 3 and 4). This further corroborates the point that the only reason to postulate exceptional explanations is a basically Eurocentric, monolingual view of language that is hooked on conceptualizing transmission as a linear process of acquisition between two parties where the entire linguistic code of A is passed directly to B and ‘acquired’ *in toto*. But this view is no longer appealing in the approach developed here, which combines evolutionary views of contact language formation with construction-based models of language transmission in socially realistic contexts. In such views, speakers *reconstruct* a language through a process of selection and replication of the features available; they do not ‘acquire’ it. As we will see in the next chapter, the SLM communities show this not only in their past histories but in the present as well, engaged as they are in a process of identity realignment in which their language is reconstructed according to recent socio-political developments, in particular globalization.

## 6.4 Final remarks

In this chapter I have traced the evolution of SLM through a study of its internal and external ecologies. In the evolution of SLM, it is clear that the socio-historical dominance of Sinhala and the cultural significance of Tamil as once a lingua franca of Monsoon Asia (chapter 2), combined with the fact that these two languages show congruence in a number of grammatical aspects, have together led to a general typological shift in which a PDM-type vernacular is restructured along Lankan features. As shown in previous accounts of the ‘genesis’ of SLM, structural observation alone (or one with little historical evidence) leads to problematic explanations. For example, structure alone cannot

explain constituent order in SLM: it is predominantly verb-final, due to the congruence of Sinhala and Tamil basic word order which tends to be verb-final. This, we might postulate, leads to postpositional morphology as seen in the case system (Slomanson 2006). But this does not apply to the verbal domain, where prefixed TA markers prevail, despite the fact that verb-final structures are more common. The focus on internal ecology can however explain why the nominal domain shows such a high degree of congruence with Lankan grammar, as seen in the development of a fully-fledged case system, while the verbal domain shows a high retention of PDM features. The typological gang-up observed in the case systems of Sinhala and Tamil, and the absence of marking of semantic roles in PDM, results in an overwhelming dominance of Lankan features in the nominal domain of the TM. On the other hand, a converse situation applies in the verbal domain: Sinhala and Tamil do not fully agree in terms of TA categories and are not particularly heavy in the aspectual department, while PDM offers a set of well-developed aspect markers. Thus SLM aligns with Sinhala – the dominant adstrate – in terms of tense, and maintains ancestral aspectual categories, due to founder effects.

What this study has shown moreover is the value of an evolutionary approach, in which language use is regarded as central, and for which a TM, which includes all varieties actually present in the contact environment, is taken into serious consideration. From such an approach, we derive the realization that there is nothing ‘special’ about the development of SLM, apart from the diversity of languages involved in the contact environment, which makes it such an interesting topic of linguistic inquiry. Its development can simply be captured as a process of attrition common to minority languages in multilingual settings, with a certain retention of material due to the high vitality of the language. This type of retention, though rare, is not unknown in sociolinguistic studies. I refrain from giving this process a name, as I do not think that new categories created ad hoc are a positive contribution to contact linguistics. I am quite comfortable with implying that the structural and social processes observed here may be unique to SLM, as it may be the case that each contact scenario is unique and that any kind of generalization will ultimately always prove too simple to really capture the evolution of a new language. (There are, however, strong similarities between the evolution of SLM and what Ross (1996) calls ‘metatypy’.) Indeed this is the case in respect to the role of intermarriage in CLF; in this chapter I have strongly argued against intermarriage as the prevalent trigger of CLF in the evolution of SLM, based on historical and linguistic data. This does not mean that I regard intermarriage as irrelevant in all contexts of CLF. As will become clear in the [next chapter](#), I acknowledge that, in many cases of CLF, historical information reveals that intermarriage does indeed play a significant role. But, crucially, *CLF may occur without extensive intermarriage*, as the case of SLM has shown. This is an example of the

danger of generalizing without sufficient information of the ecology in which languages evolve; by assuming *a priori* that CLF = intermarriage, we not only end up with a problematic account of the evolution of SLM, but we also miss the important realization that contact *outside* the household and societal multi-lingualism can also be a force behind the evolution of new varieties.<sup>24</sup>

## Notes

- 1 The quote refers to Malay soldiers and the practice of gambling. It reveals, as argued in this chapter, that womenfolk of Indonesian origin (Ambon) were present in the Malay community, and that if intermarriage occurred, it involved both Sinhalese (Singulayans) and Tamil (Malabar) women.
- 2 In Raven-Hart (1953: 69).
- 3 Recall that Bazaar Malay, the *de facto* lingua franca of the trade route stretching from Southern China to Northwest India from at least the fifteenth century, is based on colloquial Malay and Javanese, as well as borrowed elements from Hokkien (see section 3.2 of this volume).
- 4 As already argued in chapter 3, it is not easy, or perhaps possible at this point in time, to establish exactly which type of lingua franca was used. While Batavian Malay, a Malay-based lingua franca, or Jakarta Malay Creole have been postulated as good candidates for a common language in the early days of the Diaspora (Hussainmiya 1987, 1990), the establishment of Jakarta Malay is unlikely until the third quarter of the nineteenth century (Grijns 1991). Evidence of Moluccan material in SLM has also been found (Adelaar 1991) which suggests that Eastern Malay dialects, which were much more widespread across the Indonesian world at that time (Hans den Besten p.c. June 2005), may have been involved in the evolution of SLM.
- 5 Since the 1930s, Sinhala and Tamil have been mediums of instructions in Sri Lanka, together with English. After independence, it became common practice for schools to offer either Sinhala or Tamil instruction, but not both as was common during British colonial rule. Due to the religious affinity between SLM people and Tamils (both Muslim), SLM children have, until recently, predominantly joined Tamil-medium schools. Yet Sinhala remains the dominant language of everyday life.
- 6 The representation of Malay data follows conventional Malay scholarship. For Sri Lanka Malay, doubled vowels indicate vowel length, and double consonants indicate geminates.
- 7 Though never extensively described, it is likely that Lankan Tamil varies from Tamil varieties spoken in Southern India. However, the Tamil that early SLM communities came into contact with would have been typical of Moor communities originally from Southern India and therefore not classic Tamil nor Lankan Tamil. Since at this time we are not yet in possession of robust data on either Lankan Tamil or Moor Tamil, I here use the most reliable sources available based on South Indian Tamil. For both Tamil and Sinhala data, it is important to be aware of the significant differences between literary and colloquial registers.
- 8 This may apply also to the Slave Island community and other more rural communities.
- 9 Exact etymology is difficult to obtain considering the diversity of Malay/Indonesian features involved in Bazaar Malay varieties.

- 10 Note that the conflation of Experiencer, Benefactive, Goal and Possession is in fact not peculiar or unique to South Asian languages, but can be seen as a universal tendency of Dative case marking, as clearly shown in Blake (1994: 145), who describes Dative as the main non-core case used to mark complements.
- 11 Nordhoff (2009) also describes the use of *hatthu* 'one' as a definiteness marker in SLM which follows Sinhala grammar.
- 12 More parallels between Animacy distinctions in SLM and Sinhala are discussed in Nordhoff (2009).
- 13 There appears to be variation in SLM varieties here: Ansaldo and Nordhoff (2009) find evidence of a hybrid system in the Kandy variety where a third tense seems to have emerged. However the argument in favour of a clear future marker is weak: it is very infrequent and more likely to be a matter of Irrealis modality (Smith and Paauw 2006: 168–9; Ansaldo and Nordhoff 2009).
- 14 However, the progressive in SLM is really marked by the grammaticalized [*ambe...dudu*] construction (see Nordhoff 2009).
- 15 This could be either a case of retention of Malayic verb-medial word order or a result of the fact that spoken Sinhala allows substantial word order variation in the VP (Silva 2003), which may trigger selection of VO-OV patterns from the matrix. This is a common observation in CLF (see chapter 7), and perhaps a general reflection of the fact that constituent order is not always as fixed as grammatical theories like to portray it.
- 16 This is corroborated in generative analysis of the SLM VP (Slomanson 2006).
- 17 'Moors' is the term used to refer to a group of people of Arab origin, partially Tamil speaking due to prolonged residence in South India before arriving in Sri Lanka.
- 18 Note also that, until 2004, a mere one-hour recording of SLM was the main source of linguistic data available.
- 19 Hussainmiya humbly adds that his "exercise is not undertaken with any formal training in linguistic techniques ..." (1987: 157–8).
- 20 A more appropriate scientific exercise might have tried to disprove the influence of Sinhala by indeed looking for it.
- 21 Also, as argued for example in Den Besten (2007), relexification in its strict form may account for some of the processes involved in restructuring, but other explanations are often needed for a complete picture of genesis.
- 22 To the extent that this may be a viable account: for strong reasons against this, see among others DeGraff 2005, Ansaldo 2007, Ansaldo, Matthews and Lim 2007, Ansaldo and Nordhoff 2009.
- 23 Rapidity does not constitute a feature of a special diachronic process, as this work, among others, does not recognize the dichotomy between 'normal' and 'special' development (see also Ansaldo, Matthews and Lim 2007). As was argued in chapter 5, speed of diffusion relates to size and type of community; in small communities diffusion is of course relatively fast.
- 24 Similar functions are found in Ambonese Malay (though there it is pre-verbal, Minde 1997: 191).

## 7 Identity alignment in Malay and Asian-Portuguese Diaspora

---

*And so to resume, science has determined that in the beginning there was only one group – in any event not more than a few. With time, others developed. Curious, that a new group, arriving on the scene, as it were, would not first survey the terrain, size up the situation, listen to the prevailing parlance. No – it emerges with its own language, its own pantheon of gods, its own universe of traditions. With relative immediacy it demonstratively underscores its own otherness.*

Ryszard Kapuściński (2004: 172)

As suggested in previous chapters (in particular [chapters 3 to 6](#)), in multilingual ecologies we may find communities that, in certain socio-historical environments, undergo a change in their linguistic repertoire through contact. I have described these environments as often involving closely knit minority/diasporic groups of mixed ethnic origins, often in the position of intercultural brokers. I have already argued (see in particular [chapters 4 and 5](#)) that language contact and contact languages are in fact quite common in human history, exotic as they may appear to the Western monolingual speaker. We should therefore find ways in which to view such ecologies as not ‘special’ or abnormal, because multilingualism is at least just as common as monolingualism in current societies.

Too often do linguists identify a speech community with a single language, even when it is clear that the community is characterized by multilingualism, and too often do we find that one particular language is exclusively linked to the cultural identity of a group. This has, for example, been the case of the Malays of Sri Lanka: regardless of the fact that members of Sri Lanka Malay (SLM) communities are at least bilingual, commonly trilingual, and at times quadrilingual, they are almost always identified only as speakers of SLM. There is another possibility: members of multilingual speech communities may have a more complex and sophisticated pattern of self-identification, namely with a multilingual repertoire. In fact, I argue below that this is exactly the case of the Malays of Sri Lanka ([section 7.1](#)) and that the process of contact language formation (CLF) in general may be understood as the outcome of an identity alignment within a multilingual context. This means that we need to revise our



understanding of terms such as ‘language shift’ and ‘language death’ when describing certain types of multilingual ecologies, in the light of language as cultural and economic capital that is transformed and reconstructed to best serve the purpose of a community (Stroud 2000; Mufwene 2004a, 2008). Notions of ‘shift’ and ‘death’ do not apply well to multilingual environments (Lim and Ansaldo 2007). In an evolutionary, usage-based approach to language such as the one pursued here, it is indeed possible that these notions may be too limited to properly describe patterns of language change in a multilingual ecology. These notions were initially developed predominantly to describe shifts in the language practices of populations that used to be monolingual: shift is in fact defined as the replacement of one language by another as the primary means of communication (Fishman 1964). In this sense, language shift and death are usually associated with negative social, political, and/or economic issues, such as the destruction of autonomous communities or the deprivation of traditional land, and this is often accompanied by the call to reverse shift (Fishman 1991) and/or to document and maintain endangered languages. While I am very sympathetic to the discourse of endangerment, it is also important to realize that, in a multilingual ecology, language shift and language death often also involve language birth, in the sense of the formation of a new language through contact (Mufwene 2004b, 2008). Therefore it may be more appropriate to look at patterns of CLF in terms of linguistic alignment, without necessarily associating it with notions of loss, whether in terms of language or identity. The notion of alignment calls for a serious consideration of the linguistic and cultural negotiations that occur in the multilingual ecologies in which CLF occurs.

## 7.1 The ecology of identity alignment

The notion of identity alignment is framed within the anthropological approach to language and socialization introduced in [chapter 5](#) as part of a non-essentialist view of language – i.e. a view in which the use of language assumes a central role – that may prove more useful in the study and reconstruction of CLF. The views expressed within the socialization paradigm build on notions of language transmission as acculturation involving social participation in the construction and representation of identity throughout life (i.e. linguistic acts of children and adults), and assume a rich environment from which to select (linguistic) features. These have all already been previewed in the discussion of the evolution of SLM, which illustrates the following important facts about the evolution of new varieties of language:

- Contact languages do not exist in a vacuum; they are in close relation with other languages spoken in their environment, which I have referred to as ‘adstrates’. This leads to widespread multilingualism as in the case of

colloquial Sinhala, Lankan Tamil, and vernacular Malay in the ecology of SLM.

- Contact languages are often based on previous experiences of trade languages, such as Bazaar Malay (BZM) in the evolution of SLM and Asian-Portuguese in Makista (see [section 7.2](#)). As already noted in Le Page and Tabouret-Keller (1985), contact languages do not necessarily derive directly from a putative lexifier language or a pre-existing pidgin ([section 4.2](#)); they may be continuations of other contact language environments. Where Asian-Portuguese is concerned, when the Portuguese arrived in India, for example, they found at least two established lingua francas, one of Arabic and the other of Malay type ([chapters 2 and 3](#)). They also brought with them Portuguese codes from West Africa so that, strictly speaking, there was no contact between Portuguese, Malay, and Arabic *per se*, but rather between varieties of these languages used for interethnic communication.
- Contact languages arise because of the evolution of a population of speakers that is negotiating a new position in the ecology in which it is situated. In other words, CLF is the product of identity alignment in a multilingual context (Lim and Ansaldo 2007). This is akin to the cultural focussing described in Le Page and Tabouret-Keller (1985) in terms of ‘acts of identity’, and goes beyond notions of ‘target language’ and ‘restricted access’ to one specific code.

In relation to these points, I propose that the processes of admixture, i.e. selection of features from different linguistic codes and recombination in a new grammar, is the natural outcome of an identity alignment that typically occurs in ethnically heterogeneous contexts. In these contexts, a newly emerged community with multiple linguistic resources is prone to realigning its identity through a cultural accommodation that involves linguistic negotiation of the codes available in its environment (Lim and Ansaldo 2007). In so doing, the community achieves both *integration within the new context* and *self-identification through cultural innovation*.

In the rest of this chapter, I look at three interrelated aspects of contact language ecology, namely the cultural, political, and linguistic alignments of multilingual communities in which new languages emerge. I briefly revisit these aspects in relation to the Malays of Sri Lanka and relate them to the Babas of Melaka and the Macanese of Macau.<sup>1</sup> In [section 7.2](#), I illustrate how the process of identity alignment translates into grammatical output in terms of admixture and typological congruence in the new languages created in the latter two communities.

## 7.2 Multiple alignments in contact settings

The Malays of Sri Lanka, the Babas of Melaka and Singapore, and the Macanese all used to share the following traits, albeit to different degrees.

- i. Power-brokerage: all the three groups were in a relatively influential position as intercultural, multilingual mediators in the ecology in which they functioned.
- ii. Internal focus: each group had a clear cultural, political, and/or linguistic identity that set them apart as a discrete group from the rest. This identity was and still is negotiable and fluid.
- iii. Endogamy: this relates to (ii), in that the community was characterized by internal cohesion; endogamy was prevalent especially in later historical phases. I stress this factor because it is just as important in CLF as its counterpart, exogamy, which characterizes the formative stages of many new communities, and which usually receives much more attention. As we will see in this section, while Baba Malay and Makista have their origins in patterns of exogamy, the actual focussing that brings about a new identity and CLF occurs once endogamy is widespread.
- iv. Plurilingualism: the linguistic ecology of the communities being described here was a diverse, heterogeneous one, characterized by social and individual multilingualism and language and register variation.
- v. Cultural hybridity: language is only one aspect of a pluralistic society where we find variation and fusion of elements originally belonging to different systems. Other aspects of culture such as beliefs, attire, and cuisine likewise reflect the new identity alignment of these communities.

### 7.2.1 *Pluralism and its advantages*

Throughout the history of the Malays of Sri Lanka, the Sri Lanka Malay vernacular was not the sole characteristic of their linguistic repertoire. As already suggested in the [previous chapter](#), to this day the most significant aspect of the community is their multilingualism: the SLM community has always been perhaps the most multilingual of all the communities co-existing in Sri Lanka – Sinhalese, Tamils, Burghers – having in their repertoire the main languages spoken on the island, viz. Sinhala and Tamil, alongside SLM (Lim and Ansaldo 2007; Ansaldo 2008). In some social strata, the languages of the colonizers are also preserved: some older Malays speak Dutch and older and middle generations often have a knowledge of English. This multilingual repertoire put them in an ideal position as cultural brokers within Sri Lankan society, as they were able to negotiate their social position within Lankan networks as well as between Lankan society and colonial structures. Today the Malays are still clearly multilingual: 66 per cent speak at least four languages: SLM, English, Sinhala, and Tamil; and the remaining 34 per cent merely have Tamil absent from this multilingual repertoire.<sup>2</sup> Crucially, the Malays of Sri Lanka identify themselves as Sri Lankans *and* Muslim Malays (of Indonesian origin): there is no conflict in this position nor separation of traits.

Multilingualism was crucial to the Malay community as it positioned them in a powerful role between the colonial powers and the local Sinhalese and Tamil communities. Under the Dutch, people of Malay and especially Indonesian provenance were used in Sri Lanka as soldiers, guards, and servants. The Dutch were familiar with contact Malay varieties from their experience in Java and had high respect for the fighting abilities of Javanese warriors. Based on this experience, under the British, the Malay Rifle Regiment was instituted, a prestigious division used to fight local (Sri Lankan) armies. Retired soldiers and officers of the regiment were later in a very good position to become police officers, supervisors, and trusted guards in the British colonial structure. Therefore, in both Dutch and British Ceylon, many Sri Lanka Malays were in privileged positions, intermediaries between colonizers and locals, as soldiers, officers, cinnamon garden supervisors, and tea plantation managers, and were proficient in all the languages needed to interact with all parties concerned (Lim and Ansaldo 2007). At the same time, as mentioned in [chapter 6](#), the community was strongly endogamous until very recently: it is only in the youngest generation that some mixed marriages have started taking place.

Let us now turn to the Peranakan or Baba communities of Melaka and Singapore. Recall the historical picture of Monsoon Asia sketched in [chapters 2 and 3](#), in which Sino-Javanese contacts emerged as one of the dominant patterns of cultural and linguistic interaction ([section 3.3.1](#)). As a consequence of widespread intermarriages, as early as the sixteenth century the communities known as *Ioaos* (or Javaans) emerged in Java, i.e. locally born people of mixed Chinese-Javanese descent, who were recognized by the Dutch as a group which was neither Malay nor Chinese, traced their origins to China, but looked almost like locals ([sections 2.2 and 3.3](#)).

The first clear attestations of Baba family records are from the seventeenth century, i.e. during colonial rule of parts of Java by the Dutch and roughly two to three centuries after the likely origins of these communities. The first Western description of such a community comes from Crawford (1856: 96), who wrote that

the settlers, whenever it is in their power, form connections with the native women of the country; and hence has arisen a mixed race, numerous in the older settlements, known to the Malays under name of Peranakan China [i.e. *cina* 'Chinese'].

The politics of segregation introduced by the Dutch (Reid 2000) and continued by the British, under which individuals of mixed origins were used as middlemen, merchants, and interpreters between the colonial administration and local populations, lie at the very basis of the establishment of the group, characterized by a pluralistic Chinese-Malay/Indonesian identity. Under the British, the Babas became intermediaries between the colonial administration and newly arrived Chinese migrants. The cultural pluralism of these communities, and

their long-term commercial engagement in the region, gave them a clear political and cultural advantage over other groups, and made them ideal power-brokers in the ecology they inhabited.

Through their increasingly endogamous practice (section 3.3.1), by the early nineteenth century these mixed populations became culturally quite distinct and were identified as *Babas* or *Peranakan China* (the latter already seen in the quote above); today they are also sometimes referred to as *Straits(-born) Chinese*. The Babas and Nyonyas (see chapter 1, fn. 3) are very proud of their heritage and regard themselves as true indigenous groups of the Straits region; they do not identify with either Chinese or Malay groups. During British colonial rule there was, in some generations and higher social strata, a tendency to claim British identity (Ansaldò, Lim and Mufwene 2007; Lim 2009b).

The origins of the very terms with which these communities identify themselves are revealing as to their complex, shifting position at the crossroads of identities. The term *Baba* has disputed origins but is most likely a word of Middle Eastern origin that confers a title equivalent to *Sir* or *Mister* (Rudolph 1998: 38), i.e. a relatively high social standing, and was introduced to Southeast Asia through India (Tan 1988b). It refers to the male individuals born out of the union between overseas Chinese merchants and local Malay/Indonesian women, and it is also used to distinguish their particular language, Baba Malay. The Malay term *Peranakan* (in which *anak* denotes ‘child’) also indicates a group of mixed racial origin with the connotation of ‘locally born’, i.e. in the Strait of Melaka (Rudolph 1998: 42). But *Peranakan* has also become a common term for Indonesian-Chinese people, whether of recent or ancient arrival, independent of language use.<sup>3</sup> While the use of the term *Ioas* may at times imply generic Javanese origin without specific reference to ethnic Chinese influence, the use of *Babas* clearly identifies communities of Chinese-Malay/Indonesian descent that stand apart from local Malay/Indonesian as well as from (recent) Chinese immigrant groups.

A third context in which multilingualism, power-brokering and new identity can be found is the case of Macau. Recall that the Portuguese were always in a significant numerical minority in their settlements, that they took a relatively liberal stand on intermarriage and concubinage, and that they relied heavily on local labour (see sections 2.5 and 3.4). This also applied to mercantile environments (i.e. shipping), as the majority of crews were composed of Asians and Africans with a minority of Portuguese officers (Boxer 1973; Baxter 1996: 300).<sup>4</sup> Considering the historical origins of the population of Macau, deeply connected to other parts of the Portuguese Empire, in particular to Melaka, it has been suggested that the native-born Macanese originated from three main strands: Portuguese, Asian-Portuguese, and, later, (Southern) Chinese (Coates 1978: 45). In recognition of the special identity of the Macanese, the Cantonese referred to them as *tóusāang* ‘born on the soil’, and in Portuguese they are *filhos de Macau*

‘sons of Macau’. Da Silva (2001) identifies them as people who are (i) locally born, (ii) of Portuguese descent, and (iii) with Malay or Chinese mothers. Besides heritage, Macanese identity is also clearly defined by (a) the Christian faith, (b) adoption of Chinese customs, and (c) political separation from Portugal.

The strong Christian faith that set the Portuguese apart from neighbouring groups has also been noted to have contributed to a certain social impermeability observed for other Asian-Portuguese communities, e.g. in Melaka (Holm 1989: 292). The theme of isolation from Portugal too recurs often in Macanese historiography and is particularly significant (da Silva 2001: 155). Macau’s separation from Portugal was not only geographical in the sense of remoteness but also political and economic; this was related to the fact that in 1580 Portugal virtually lost its independence as a consequence of the union of Portuguese and Spanish crowns under the Castilian king of Spain.<sup>5</sup> Macau never acknowledged this transition of power and remained fiercely independent, neither being Portuguese nor accepting Spanish rule (da Silva 2001). It was crucial to Macau to keep this new political status secret in the eyes of China, for fear that Chinese authorities might take advantage of the situation and undermine its independence. From a Macanese perspective, Portugal felt remote because it was now a defeated country that appeared poor and backward compared to the opulence of Macau. Thus, only half a century after its establishment, Macau was already a very independent colony, connected to Melaka and Goa, but also in competition with them. From then on, the history of the Macanese would be a history of adjustments, concessions, negotiations, and adaptations to their new immediate environment.

In the Macanese community, multilingualism in Portuguese, Makista, and Cantonese was common (Amaro 1994: 219; Teixeira 1994). Through their knowledge of the Chinese languages as well as their understanding of Chinese customs, the Macanese acted as mediators between other Western traders and Chinese authorities in the Canton Trade (see section 8.2). These advantages allowed them a privileged position with the Chinese who used them as a buffer between themselves and the other ‘barbarians’ knocking at their doors until as late as the mid-nineteenth century. The Macanese were also respected by Western traders as the only ‘Westerners’ allowed to trade directly with Japan, which confirmed their superiority in intercultural communication. Even if in a demographic minority, the Macanese became power brokers between China and the West, a position that allowed them to determine their own destinies and strongly influence their own geopolitical environment throughout colonial history (see chapter 8).

### 7.2.2 *Cultural fusion*

The cultural pluralism that characterizes the communities described here can be seen in novel recombinations of cultural traits drawn from the diverse ecology

in which they evolved. As noted in [chapter 6](#), the Malays of Sri Lanka have acculturated to various aspects of South Asia, such as in their attire. They have been wearing the *sari* and *shalwaar kameez* instead of the Malay/Indonesian *sarong kebaya*, which has appeared in their wardrobes only in the last few decades. While they have maintained their original Muslim faith, they have adopted a Hindu wedding practice that consists of payment of a dowry, in contrast to the Islamic tradition of the groom paying *mahar* to the bride's father (Saldin 2003; Lim and Ansaldo 2007). In their cuisine they combine elements of Malay/Indonesian origins with traditional South Indian dishes.

Non-linguistic examples of the unique, hybrid culture that characterizes Baba communities include a mixed cuisine largely influenced by Malay, Indian, and Javanese elements. The *nyonya* cuisine is one of the earliest and most successful examples of 'fusion' and is very revealing as to the diverse cultural origins of the Babas. One of its typical dishes is *Babi Buah Keluak*, a dish of pork (ribs) cooked in tamarind juice with large black nuts (*buah keluak*). These nuts, which have a hard shell and a meaty, dark pulp that is dug out with the help of a little spoon, only grow in Java and Sumatra. Other Indonesian influences are found in *Sambal Goreng Java*, a dish of prawns and eggs in coconut gravy originally from Java as its name suggests, and *Beef Rendang*, a spicy beef casserole with coconut and tamarind typical of Sumatra. This confirms that, though well-known Baba communities developed on the Malay Peninsula (Melaka and Penang and later Singapore), the origins of the women were most often in Indonesia, as the Babas predominantly married servants, slaves, concubines, and other non-Muslim Malay women. Chinese influence is also found in the numerous *Hokkien Mee* (Chinese noodles) dishes that feature in *nyonya* cuisine, and *Mee Siam* which indicates a Thai connection,<sup>6</sup> as well as Malaysian ones, for example, in the various *Laksa* recipes, noodles in spicy coconut soup often featuring seafood.

Besides cuisine, attire also reveals cultural admixture, with the wearing of Malay/Indonesian *sarong* and *kebaya* (instead of the Chinese dress) by the Nyonyas. At the same time there is retention of Chinese rituals, such as traditional wedding customs involving imperial-era wedding costumes (Tan 1988b: 299). Thus, it is important to note that language is not unique in showing mixed traits of different origins, but that other socio-cultural aspects are similarly hybridized. It is typically under such circumstances, i.e. formation of what is perceived of as a new culture, that a new language emerges.

In the Macanese sphere, examples of cultural admixture are likewise found in attire, cuisine, and even beliefs. While the Macanese adhere to the Christian Catholic faith, they are strong believers in *feng shui*, the Chinese practice of geomancy, according to which the structure and content of the space we inhabit can influence our destiny (Amaro 1994).

Macanese cuisine is particularly revealing in the use it makes of spices typical of South (and Southeast) Asia, such as saffron, cinnamon, curry, tamarind, and



turmeric, as these indicate a cultural connection to India and Southeast Asia. For example, fish chutney, which today can be based on *bacalhau* (cod), features prominently, revealing Indian influence. Typical ingredients also include glutinous rice, soy sauce, and Chinese sausage, used in dishes such as the *Arroz Lap Mei*, glutinous rice cooked with Chinese meats (pork/liver sausage) and cuttlefish, revealing the Chinese influence in Macanese cuisine. Bakery and pastry clearly manifest Portuguese influence, as dough and egg feature much more prominently among Macanese desserts than the traditionally lighter, soupy Asian sweets. At the same time, the ingredients of desserts are often of Asian origin, such as sago, coconut, and jaggery. Macanese dishes include the use of spice pastes, such as *balichão*, a sauce made by grinding small dried shrimps and mixing it with chillies and salt, typical of Malay cuisine (also found in *nyonya* cuisine as *belachan*). This is used to prepare the *Balichão Caranguejo com Flores de Papaya*, a spicy crab dish with banana flowers, one of Macau's favourite dishes. Dishes such as *Furusu* combine a Japanese etymology with typical Portuguese ingredients: it consists of cold meats dressed in ginger, chillies, mustard, and mint. The study of Macanese dishes is extremely fascinating and complex also because it is not always easy to identify clearly the origins of a recipe: *Aluar*, for example, is one of the most typical desserts, based on almonds, very common in Spain and Portugal, but is originally of Arab origin and probably reached Macau via the Iberian peninsula (Amaro 1994: 44).

Regarding attire, it appears that Macanese males dressed in European fashion, perhaps with lighter jackets, and dispensing with socks and boots. In their homes, however, they often wore *calça moura* and *baniane*, loose trousers and long shirt of Moorish style, as in Goa. Women followed Indo-Malay style, probably also under the influence of Goa (Amaro 1994: 50). They wore the *baju*, which consists of a piece of cloth wrapped around the waist like a Malay *sarong*, a light shirt similar to a *kimono* and a *saraça*, a piece of cloth used to cover the head while outdoors. For festive occasions they often wore a long *cabaia*, which is in fact a short dress usually of silk, with rich embroidery and high slit on the sides, similar to the Chinese *cheongsam*.

Other aspects of cultural diversity typical of the Macanese can be found in the passion for the game of *mancala*, an African chess-like game that had become extremely popular in Melaka (known there as *chongkak*). Of Macanese creation was the game of *bafá*, a type of domino adapted from a Chinese game. Asian influences are also found in the type of embroidery and needlework that Macanese women produced. Finally, Indo-Portuguese influences can be found in typical Macanese lullabies and songs, a number of which show similarity to those recorded in Daman (Dalgado 1917; Amaro 1994: 54). Curiously, the *Travesti*, a musical performed exclusively by men wearing female clothes, is a typical feature of Malay culture found in Macau and other mixed ethnic groups with cultural links to Melaka such as the Babas.



### 7.2.3 *Identity alignment and language evolution*

#### 7.2.3.1 *Sri Lanka Malay (SLM)*

The evolution of SLM itself can be seen as a process of interacting and of negotiating linguistic identities in a new environment. These interactions and negotiations are still ongoing, albeit with different linguistic codes involved, due to economic, cultural, and political changes. This can be seen in the growing importance of the use of English in the household, at times replacing the SLM vernacular, and the recent addition of *Bahasa Melayu*, the standard variety of Malaysia, to their repertoire, which may eventually replace SLM.

The notion of *identity alignment* argued for in the case of SLM in Lim and Ansaldo (2007) implies two significant things. First, the implication is that identity can be defined by being multilingual. This requires a loose view of ‘community’ in which identities can be adapted and renegotiated in an ever-changing ecology (Djité 2006: 12). The idea of simultaneous identities (Woolard 1999: 20–1) in this context suggests that, instead of assigning one language a preferential status in a multilingual community, one’s linguistic identity is shaped by the plurality of linguistic codes itself, regardless of the order in which they were acquired. In the case of the SLM community, we clearly see a multiple linguistic and cultural identity that is both Lankan *and* Malay. Second, as observed by Le Page and Tabouret-Keller, concepts such as ‘native language’ and ‘mother tongue’ mean little in multilingual settings, where more than one language is spoken from birth (1985: 189). Ferguson (1982: vii) suggests that we “quietly drop” the notion of mother tongue from professional linguistics, a call heeded by Rampton (1990: 107), who argues that the related notions of expertise, inheritance, and affiliation do not adequately identify sociolinguistic situations and that it is inaccurate to view people as belonging to only one social group. Culturally, Sri Lanka Malays are comfortable with being identified as Sri Lankans and Malays. The former identification is manifest in the absence of a wish to ‘return’ to Malaysia or Indonesia, while the latter is expressed in their interest in revitalizing the Malay language in the community.

In SLM communities such as Kirinda (see chapter 6), SLM may still be the primary language (especially in the home domain), but it is far from exclusive: code-mixing with Sinhala and/or Tamil is frequent and both these languages enter the lives of Kirindians from very early on, through play, school, and media. Sinhala and Tamil also occupy very significant functional roles in the community, as languages of education and economic opportunity. In other words, Sri Lanka Malays can be seen as having a pluralistic identity, and are not conflating or changing identities; rather, this identity is simply multifaceted, as is their linguistic repertoire. In this sense, the linguistic profile presented in chapter 6 falls short of describing a variety of SLM: for such a

goal to be achieved, we do not need a ‘grammar’ of SLM, but rather a comprehensive description of trilingual admixture that defines the communicative practices of the community. The lack of a need to rigidly associate Sri Lanka Malay identity with the Sri Lanka Malay language is markedly observed especially in the views of the younger generation, who, in response to the statement ‘Speaking SLM is crucial to my identity as SLM’, chose the answer ‘strongly disagree’.<sup>7</sup> Moreover, this fluid identity has not changed. In colonial times, the Malays had within their multilingual repertoire SLM, Sinhala, Tamil, English, and these different linguistic varieties served various functions and allowed various identity alignments.

### 7.2.3.2 *Baba Malay (BM)*

Baba Malay, the language of the Babas, shows negligible Western influence until the nineteenth century. This is important because it corroborates the historical evidence that BM varieties were already developed before the arrival of Europeans in the sixteenth century. As an already stable language variety combining Hokkien and Malay/Indonesian features, it was less influenced than varieties that developed in ecologies where Western languages were present. Until the nineteenth century, the linguistic ecology of the Babas included BM, Hokkien, and Malay; the ties to Fujian were strong, maintaining Sinitic varieties in the ecology of the Babas, and in the immediate environment of the Baba and Nyonyas were Malay/Indonesian vernaculars and contact Malay varieties spoken in the Malay peninsula and on the northern coast of Java (see [sections 3.2 and 3.3](#)). BM represented an ancestral vernacular of political and economic advantage that afforded the Babas a position as cultural brokers in their environment. At the same time, it distinguished them from local Malay/Indonesian groups as well as new migrant groups from China.

During the nineteenth century, a new element entered the ecology, as English became an increasingly important language in Southeast Asia, and the Babas were among the privileged few who acquired and benefited from it (Pan 2000: 172). Being wealthy merchants of high social standing, the Babas held a high regard for English-medium education and sent their children to English-medium schools, which they were also instrumental in establishing (Tan 1988a: 52). By the mid-nineteenth century, their ability to converse in this colonial language had further strengthened their prominent socio-economic position with respect to other local communities, to the point where they were in fact sometimes referred to as “King’s Chinese” (Tan 1988a: 53). The growing importance of English among them as a lingua franca, as a vernacular, and as an identity marker can be clearly seen in the debates that arose during the mid-twentieth century. Though there is no doubt that most Babas identified themselves with BM (Tan 1988b), around this time more and more Babas were also aligning themselves with English. While BM had competed fairly well

with English as a lingua franca until the early twentieth century, by the end of the 1960s English almost completely prevailed as the means of interethnic communication *par excellence* (Rudolph 1998: 335). The younger generations of Babas and Nyonyas in Singapore, for example, have at best a passive knowledge of BM, and were often raised in English-speaking homes (Lim 2009b). In other words, the Babas have adjusted to a new ecology in which BM no longer affords them a cultural advantage; moreover, they align themselves at present with an ecology that is increasingly less multilingual, due to the rise of standard languages and the policies of normative usage advocated in societies such as Singapore (Lim 2007, forthcoming; section 3.3.3 of this volume). The strong influence of British culture, observed in the alignment to the English language, which the Babas also contributed in spreading in their society, shows a later stage of identity alignment in the evolution of Baba communities (Ansaldò, Lim and Mufwene 2007; Lim 2009b). As we have seen in the case of the SLM communities, recurring linguistic alignment is particularly salient in communities characterized by linguistic and cultural admixture; this is a consequence of their multilingual practices, which allow them a wide spectrum of linguistic negotiation, as well as a historical heritage of their socio-historical conditions of minority-migrant groups adept at constantly adjusting to changing environments.

### 7.2.3.3 *Makista*

Makista, the Macanese vernacular – or *patoá* as it is often referred to in the literature on Macau – is a form of Asian-Portuguese (Baxter 1996) with Sinitic elements (Ansaldò and Matthews 2004). Macanese literati identify the early stages of their vernacular as a colonial dialect, uneducated, and imperfect, but nonetheless the mother tongue of the early Macanese around the mid-sixteenth century (Batalha 1994; da Silva 2001). It consisted of Portuguese mixed with Chinese as well as some Canarim,<sup>8</sup> Malay, and “a drop” of Spanish (Ferreira 1983). It is clear that whatever dialectal variety of Portuguese was used in Macau was limited and was most certainly a non-standard, colloquial variety rich in foreign influences, typical of sailors who participated in the African, Indian, and Asian trades (Arana-Ward 1977).

In the sixteenth century there was a dramatic influx of Chinese people fleeing to Macau as a result of the Manchu invasion of Canton 1652. These migrations of mostly Cantonese and Min speakers continued in increasing numbers until the middle of the nineteenth century. Chinese–Macanese intermarriages and Chinese conversions to Christianity are reported in the records of marriages and christenings from the parishes of Sé and St. Lourenço (Teixeira 1994: 119–27). Throughout this phase, *patoá* was spoken by everyone, whether Macanese or Chinese, male or female, whether in the market or in church (da Silva 2001). Makista texts from this period are written in a vernacular with

Asian-Portuguese features such as tense, mood, and aspect (TMA), but also verbal tense morphology and reduplication. This suggests register variation and possibly a coexistence of a more literary and a more colloquial variety. We can thus identify a process of alignment in which Sinitic features are added to the Malayo-Portuguese base, probably between the seventeenth and the eighteenth century, as a result of an increased multiethnic environment and established political and cultural autonomy.

The middle of the nineteenth century marks a period of further change, with schooling in English and standard Portuguese entering the ecology. Indeed, in the twentieth century, Makista was said to be ‘decreolizing’ towards Portuguese (Baxter 1996). This was however also a period of cultural indeterminacy due to the fact that, after the restoration of democracy, Portugal entered a rapid decolonization process. It became apparent that Macau would return to China and the date was fixed for the official transition to happen in 1999. This was also a period of modernization under the influence of the British colony of Hong Kong. As a consequence of these trends, the last few decades saw a substantial population movement from Macau to Hong Kong and Portugal. While European Portuguese gained popularity in Macau, the proximity of Hong Kong and the financial advantage of speaking English resulted in limiting the alignment with Portuguese: English entered the multilingual ecology of Macau, reflecting a general trend in postcolonial Asian cities. This last phase has seen a dramatic decrease in the usage of Makista, partly due to convergence towards standard Portuguese (Tomás 1990: 63). In the words of one of the most prominent literati of Macau, known as ‘Adé’ (José dos Santos Ferreira 1978: 6): “In the first decades of this century, the Macanese dialect – by then commonly known as ‘patois’ – was still spoken among several Macau families ... From the ’30s onwards the dialect fell into disuse in Macau.” It seems however that *patoá* is still spoken in Macau today, albeit only in some isolated families and perhaps only by senior citizens (da Silva 2001: 170; Alan Baxter p.c. 2007). However, Macanese identity is still strong, as indicated by the cultural vitality found in Macau as well as Macanese Diaspora across the globe.

### 7.3 Identity alignment and admixture

The linguistic outcome of identity alignment is selected grammatical recombination of existing features of languages into a new grammar; this does not however mean that the other linguistic codes are abandoned, but that a new one is created that uniquely defines the new cultural profile that emerges through identity alignment. As we have seen in the case of SLM, when typological congruence is present, it becomes the driving force in the formation of a new grammar. That is to say, the more frequent features of the typological matrix (TM) – in the sense of being statistically and typologically dominant – tend to

be selected in the restructuring process. Likewise, the particular vernaculars that characterize the Babas and the Macanese show a high retention of the congruent features of the linguistic codes available in their ecology. From the point of view of the speakers of the ancestral language, admixture means a combination of elements coming from different extant languages. This combination makes perfect sense in the light of identity alignment, and can be seen as involving two related linguistic acts: adaptation to the immediate linguistic norms *and* maintenance of linguistic identity markers. This suggests that, in contact ecologies, accommodation and divergence may be complementary processes rather than mutually exclusive ones.

In this section I look at the structural profiles of Baba Malay and Macanese *patoá* and compare their grammars with the significant linguistic codes in their ecologies. For BM, this means looking at other contact Malay features (or PDM) as well as the influence of Hokkien. For Makista, it means looking at Asian-Portuguese features as well as the Sinitic element. I specifically focus on non-standard vernacular data when discussing the TM, because it is quite clear that standard or literary varieties of Malay were not involved in the evolution of BM, neither was metropolitan Portuguese involved in the formation of Makista (chapters 2 and 3). This is crucial for a methodologically sound discussion, as will become clear below.

### 7.3.1 *Identifying the matrix*

The history of contact Malay varieties and the social dynamics that characterize them have already been outlined in chapters 2 and 3. Based on the reflections presented there, the TM in which BM evolves could potentially contain the following components: (i) Bazaar Malay (BZM) or Pidgin-derived Malay (PDM) or lingua franca features of Malay/Indonesian origin; (ii) Malay-Indonesian vernaculars; (iii) Hokkien (and probably other southern Sinitic varieties); and (iv) colonial languages: Portuguese, Dutch, and English. I suggest that, in looking at the evolution of BM, the predominant linguistic encounters were between (i) to (iii) above, namely a Bazaar Malay variety, probably used by Chinese merchants in their everyday environment, Malay/Indonesian vernaculars spoken by the women, and *some* Hokkien spoken by the males.

I stress ‘some’ when referring to Hokkien because the role of Sinitic has been somewhat exaggerated in accounts of BM (and BZM) so far. There is evidence that the Chinese traders who married Indonesian women in Java, for example, often did not use Chinese in the home with their children (and wives) but resorted instead to BZM (Dreyfuss and Oka 1979). And in point of fact, there are varieties of Indonesian-Chinese that combine Javanese grammar with Indonesian Malay content words and hardly show any trace of Chinese. In the case of the Babas of Melaka, though, it is clear that the Sinitic element was

stronger, particularly in the Penang community that today speaks Hokkien. However, as I will discuss in this section, it should not be overstated.<sup>9</sup>

The fact that features of BM and Malay/Indonesian colloquial registers would have been the most frequent ones in the ecology of Baba communities is fundamental from a methodological point of view, as it renders approaches that take the full system of Hokkien or literary Malay varieties as terms of comparison highly questionable: asking why a language like BM does not have Hokkien tones or an elaborate system of classifiers (e.g. McWhorter 2007: 231) is vacuous in this sense, as the answer is simply ‘because they were not prominent in the TM of the language in the first place’.<sup>10</sup> As already discussed in [chapter 3](#), other Malay registers, such as literary Malay or conservative indigenous Malay/Indonesian dialects, would have been largely absent from the ecology of BM and should not be expected to have left much trace. It is indeed most likely that Javanese and Sundanese played a more prominent role than Indonesian Malay, as the women were mostly from Java, where these two languages are more dominant than Indonesian Malay varieties; this renders a comparison to Malay features also problematic. Consequently, whether ‘pure’ Hokkien was spoken by the Babas or whether congruence of Sinitic traits was already taking place at the same time cannot be established.

In order to have a realistic point of comparison, in reviewing the basic features of BM, I compare them with (a) BZM as spoken in Singapore today (Khin Khin Aye 2005), which is as close as we can get to a specific BZM variety as well as to the language that the Chinese traders would have used in the community; and (b) Cocos Malay, a Malay/Indonesian colloquial variety whose origins include a Javanese-Malay element (Adelaar 1996b; Ansaldo 2002), and as such contains features of the Javanese-Malay type that would have been used by many of the women who contributed to the ecology of BM. Portuguese and Dutch hardly feature in BM. The absence of the former is explained by the time-depth of BM, already formed before the arrival of Western powers. The absence of the latter is likewise to be expected, even considering that the Dutch ruled over the Indonesian region for a long time. As noted in [chapters 2 and 3](#), the Dutch left few linguistic traces of their sojourn in the region, a result of their relying heavily instead on Malay and Asian-Portuguese varieties in their interactions with local powers.<sup>11</sup> Conversely, the lack of influence of these varieties in the TM of the Babas corroborates the observation that BM must have already been established before the arrival of the Portuguese in the fifteenth century.

Malay and Sinitic also play a relevant role in the formation of Makista. From the socio-historical sketch presented in [section 7.2](#), we must assume that the TM in which Makista evolved initially included vernacular Portuguese, Malayo-Portuguese as spoken in Melaka, and possibly some Indo-Portuguese elements (Ansaldo and Matthews 2004). In a later phase, Sinitic material, as well as

sporadic elements from other languages such as Spanish and Japanese, entered the linguistic ecology of Macau. It is important to bear in mind that, among all the colonies Macau had contacts with, Melaka and Goa were by far the most influential. The ties between Melaka and Macau appear particularly strong, as often remarked upon in Macanese historiography, probably due to the triangular trade between Melaka, Macau, and Nagasaki, in which spices and other products from Southeast Asia and China were exchanged for silver in Japan.

In what follows, the emphasis lies on the contact features of Malay varieties (or PDM features, see [chapter 3](#)), as seen in BM, Singapore BZM, and Cocos Malay (CM), with some reflections on Sinitic elements in Makista.

### 7.3.2 *Phonology and lexicon*<sup>12</sup>

Unlike in other aspects of grammar, as will be seen in [section 7.3.3](#), BM conforms to Malay in phonology (Lim 1988: 14); this can also be seen in the resemblance to Cocos Malay phonology, already noted by Lapsley (1983: 5) and confirmed by Adelaar (1996b: 181). BM consonants follow Malay, including a rich set of palatal consonants which notably are absent from Hokkien; the few exceptions to this congruence are found in consonants that are not indigenous to the Malay system but were introduced through lexical borrowing from Arabic at later stages of development (Lim 1988: 15). Vowels are identical in BM and Malay, with an additional schwa found in urban BM. Like Malay, BM has no tones, while Hokkien has a system of contour tones, three registers and a complex tone sandhi. Loanwords also occur without tones and are systematically modified to adapt to the BM system.

Singapore BZM, on the other hand, shows phonological influence depending on the dominant language of the users. Sinitic influence is noticeable in speakers of ethnic Chinese origin (Khin Khin Aye 2005) and Indian influences can be found in speakers of South Asian provenance (Sasi Rekha 2007). As observed in [chapter 3](#), phonological variation between speakers is very high in these varieties (see also Collins 1984 for Malaysian BZM and Lim 1988 for Melaka BZM), resulting in a complex phonemic inventory (Khin Khin Aye 2005). Such a contrast has been taken to suggest that BM is a more stable, focussed register, in which a homogeneous phonological system has evolved within the community, while BZM varieties, on the other hand, being spoken within more diffuse networks, lack this stability (Lim 1988).

On the other hand, the lexicon of BM shows significant Hokkien influence in words pertaining to salient cultural areas such as clothing, cuisine, rituals, and religious and other abstract Chinese concepts (Pakir 1986; Tan 1988a). Examples include *popiā* ‘Chinese spring roll’; *mi* ‘noodle’; *konsi* ‘business’; *ho miā* ‘fortunate’; *u hau* ‘filial’; *seng* ‘spoilt’ (Pakir 1986; Ansaldo and Matthews 1999).<sup>13</sup>



In Cocos Malay (CM), lexical admixture is readily notable: some examples, taken from Adelaar (1996b), show evidence of a variety of linguistic sources. Of Javanese origin we find *angon* ‘to look after’ (Javanese ‘to herd’); *cucut* ‘shark’; *gete* ‘raft’; *maras* ‘lungs’ (in Javanese, of animals); *walikat* ‘shoulderblade’. From English we have *gavman* ‘government’; *dek* ‘floor’; *oda* ‘order’; *opis* ‘office’; *soldia* ‘soldier’; *sto* ‘fishing stock’ (either English or Dutch origin). From Dutch, items include *kontan* ‘to pay cash’ (Dutch *kontant*); *prop* ‘plug’ (on a bottle); *skoci* ‘boat’ (Dutch *schuitje* ‘little boat’); *snapan* ‘gun’ (Dutch *snaphaan*, a type of gun). And from Portuguese there is *biola* ‘violin’ (Portuguese *viola*); *martilo* ‘hammer’ (Portuguese *martelo*); *por* ‘for’, a Benefactive marker.<sup>14</sup>

In the Makista lexicon, admixture of Asian-Portuguese, Malay/Indonesian and southern Sinitic elements is clearly reflected (Batalha 1977, 1994; Fernandes and Baxter 2001; Ansaldo and Matthews 2004). Examples from Malay (via Melaka) include: *capir* ‘to clasp’ from Malay *kapis* (*capi rupa* ‘wooden eggs’); *saiám* ‘pain’ from Malay *saying*; *estrica* ‘pressing iron’ from Papia Kristang *istrika*. From Sinitic, specifically Yue and Min varieties of Chinese, examples include: *agucham* ‘fishing net’ from Makista *agu* + Cantonese *hātsāng*; *tancá* ‘boat’ from Cantonese *dāangaah* ‘boat people’; *choncar* ‘to bump’ from Cantonese *johng* ‘bump’ + the Portuguese verbal suffix *-ar*; *faichis* ‘chopsticks’ from Cantonese *faaijǐ* ‘chopsticks’. Indo-Portuguese influence is also found, e.g. *apa* from Dravidian *appam*, a steamed flour pancake; *chareta* from Tamil *chiratta*, a type of ladle. According to Batalha (1977), the Makista lexicon contains at least 17 per cent of Sinitic words, next to 35 per cent of Malay items. As shown in chapter 8, Makista clearly bears some Sinitic influence, though this may have occurred late in its development.<sup>15</sup>

The most interesting area of lexical influence in Malay contact varieties is found in the elaborate kinship systems as well as the pronominal forms of address. Not only do these reveal different patterns of recombination of Malay and Sinitic features, but also, and perhaps more importantly, they reflect cultural admixture of Malay and Chinese practices. Table 7.1 illustrates the more frequent kinship terms used in BM (collated from Tan 1988a: 210ff. and Pakir 1986: 106–7) and their Hokkien and Malay equivalents.

It is noteworthy that elaborate systems of address are also a common feature of Javanese varieties (Aone van Engelenhoven p.c. October 2007), as we will see for CM. In this sense, the presence of kinship terminology *per se* is not an influence uniquely attributable to Sinitic; it should be taken to indicate congruence of cultural practices that emerge in the new language. Compare the CM terms of address listed in table 7.2 with those in table 7.1: the Malay origin of many terms is obvious.

Note that in CM, as in Malay varieties in general, kin terms are more widely used than proper names or pronouns in everyday discourse. This may explain why, in the pronominal system of CM – unlike in the kinship system – Hokkien elements are readily integrated next to Malay ones, as shown in



Table 7.1. *Kinship terms in BM, Hokkien, and Malay*

| Kin term                                  | BM                 | Hokkien         | Malay                                            |
|-------------------------------------------|--------------------|-----------------|--------------------------------------------------|
| elder brother/similar-aged or older male  | <i>hiā</i>         | <i>toa hiā</i>  | <i>abang</i> ('brother')/ <i>pacik</i> ('uncle') |
| elder sister/similar-aged or older female | <i>taci</i>        | <i>toa ci</i>   | <i>kak</i> ('sister')/ <i>macik</i> ('aunt')     |
| mother-in-law                             | <i>nio</i>         | <i>niu</i>      |                                                  |
| maternal grandmother                      | <i>gua ma/mama</i> | <i>gua ma</i>   | <i>nenek</i>                                     |
| paternal grandfather                      | <i>lai kong</i>    | <i>lai kong</i> | <i>nenek</i>                                     |
| father's older brother                    | <i>m-peh</i>       | <i>a-peh</i>    | <i>bapa saudara</i>                              |
| father's younger brother                  | <i>ng-cek</i>      | <i>a-cek</i>    | <i>bapa saudara</i>                              |
| mother's brother                          | <i>ng-ku</i>       | <i>a-ku</i>     | <i>bapa saudara</i>                              |

Table 7.2. *Cocos Malay terms of address*

| Usage                 | Term of address                                                                                                                                                             |
|-----------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Used by/with females  | <i>mak</i> : mothers, adults<br><i>bik</i> : adults without children/default female<br><i>(ka)kak</i> : older sister/higher status female (same generation)                 |
| Used by/with males:   | <i>pak</i> : fathers, adult males<br><i>nek</i> : grandfathers, seniors<br><i>man</i> : default male<br><i>(a)bang</i> : older brother/higher status male (same generation) |
| Specific for seniors: | <i>wak</i> : senior, childless                                                                                                                                              |
| Specific for juniors: | <i>(cu)cuk</i> : grandchild                                                                                                                                                 |

table 7.3. Hokkien influence can also be seen in the pronominal system of BM, which involves a first and second person singular pronoun of Hokkien origin, with second and third person plural forms in which the Hokkien *lang* 'people' appears as the Malay calque *orang*, as illustrated in table 7.3. These pronouns are one of the most robust features of PDM varieties in general.

In the case of BM, kinship terms as well as pronouns of Hokkien origin are usually restricted to the family domain, i.e. they are used for intimacy but not with outsiders, with whom Malay and English terms are used. This clear-cut distribution may be seen as an aspect of maintenance for the purpose of cultural preservation and differentiation from other groups (Thurgood 2001: 47). As shown in table 7.3, there is more variation within the pronominal system of BZM. In BZM, *wa* (from Hokkien *gua*) and *lu* indicate intimacy and friendship, while the Malay forms are regarded as more polite and imply social distance (Khin Khin Aye 2005). In CM, a similar situation holds, with *gua* and *lu* regarded as rude or very casual and the latter usually restricted to children's talk.

Table 7.3. *Pronominal systems of BM, Bazaar Malay, Cocos Malay, and Hokkien*

| Pronoun     | BM              | Bazaar Malay            | Cocos Malay          | Hokkien                                       |
|-------------|-----------------|-------------------------|----------------------|-----------------------------------------------|
| I           | <i>gua/saya</i> | <i>wa/saya/aku/</i>     | <i>saya/uk/(gua)</i> | <i>gua</i>                                    |
| you (sing.) | <i>lu</i>       | <i>lu/awak/engkau</i>   | <i>ma/pa/(lu)</i>    | <i>lu</i>                                     |
| he/she/it   | <i>dia</i>      | <i>dia</i>              | <i>mia/prempuan</i>  | <i>dia</i>                                    |
| we          | <i>kita</i>     | <i>kita/kitaorang</i>   | <i>kita</i>          | <i>lun lang (incl.)/<br/>lan lang (excl.)</i> |
| you (plu.)  | <i>luorang</i>  | <i>luorang/lu semua</i> | <i>do(ra)ng</i>      | <i>lin lang</i>                               |
| they        | <i>diaorang</i> | <i>diorang</i>          | <i>dong</i>          | <i>in lang</i>                                |

### 7.3.3 Grammar

Typical features of BM grammar noted in most descriptions of this variety are listed below (for a full account, see Lim 1988; also Pakir 1986 and Ansaldo and Matthews 1999):

- i. Topic–Comment structure
- ii. Reduplication
- iii. Word order in NP – demonstratives that precede the noun: *ini* ‘this’ and *itu* ‘that’
- iv. Zero copula
- v. Existential constructions based on Malay *ada* ‘to have’
- vi. Possessive constructions based on Malay *punya* (‘owner’ > ‘to possess’): [possessor + *punya* + possessee]
- vii. Causative constructions based on Malay *kasi* ‘to give’ and *buat* ‘to make’ preceding the verb
- viii. *Kena* passive
- ix. Isolating morphology

Features iii to vii have been listed as common features of most PDM varieties in Adelaar and Prentice (1996; section 3.2 of this volume), i.e. features possibly spread through the use of a Malay lingua franca in the region. I discuss the above features in order.

*i. Topic–Comment structure* BM is a Topic-prominent language, as noted by Lim (1988: 42), who lists at least four classic Topic–Comment patterns of Chinese languages found in BM:

#### (a) Subjects as Topic

- (1) [*Nika punya barang*] *semua ada sama lagi tak*  
wedding poss thing all emphatic-have same still not  
‘Wedding things, are they still the same?’

- (b) Objects as Topic
- (2) [*Chinese customs*] *dia pake*  
 Chinese customs      he use  
 'The Chinese customs, he observes'
- (c) Sentences as Topic
- (3) [*Gua pigi Kuala Lumpur belaja*] *lagi baik*  
 I      go KL                  study    more good  
 'For me to go study in KL is better'
- (d) Subordinate clauses as Topic
- (4) [*Dulu orang mati*] *coffin lama-lama letak ruma*  
 before people die    coffin    long-long    leave    home  
 'When people died, the coffin used to stay in the house for a long time'

Since Topic–Comment structures are a trademark of Sinitic varieties, Lim (1988) ascribes these features in BM to Hokkien influence, a conclusion also reached in Ansaldo and Matthews (1999). Khin Khin Aye (2005) finds similar parallels between topics in BZM and Sinitic, and Bao (2001) presents structural parallels between topics in Mandarin and Singapore English. Sinitic varieties are particular in that Topic-prominence is dominant independently of the grammatical role, as clearly shown in the examples above. It is also true that Topic–Comment structures are widespread in languages of East and Southeast Asia and can also be found in Malay/Indonesian languages and PDM varieties (Rafferty 1987). This suggests that selection of Topic–Comment as a prominent discourse strategy is motivated both by the prominence it has in Sinitic typology and the congruence between Sinitic and Malay/Indonesian typology.

*ii. Reduplication* Reduplication of attributive verbs, and to some extent adverbials, is productive in many contact varieties of Malay (Lim 1988; Ansaldo and Matthews 1999). The attributive verb, when reduplicated, may indicate intensification and/or repetition in BM (5 and 6). This type of reduplication is also common in BZM (7), where nominal reduplication also occurs (Collins 1984: 165; Khin Khin Aye 2005: 101), and in CM (8 and 9), though here the use of *skali* 'very' as intensifier is perhaps more common than reduplication.

- (5) *Makan pelan-pelan* (BM, Lim 1988: 32)  
 eat      slow-slow  
 'eat slowly'
- (6) *Bole beli barang banyak-banyak*  
 can buy    thing      much-much  
 'can buy things in large quantities'
- (7) *Lagi-lagi pergi pusing susah* (BZM, Khin Khin Aye 2005: 99)  
 more-more go    market difficult  
 'Going to the market is increasingly difficult'

Table 7.4. Reduplication patterns in Singlish and adstrates

| Pattern                              | Singlish        | BZM                          | Hokkien         |
|--------------------------------------|-----------------|------------------------------|-----------------|
| N+N                                  | intimacy        | plurality                    | intimacy        |
| V+V                                  | attenuation     | attenuation,<br>continuation | attenuation     |
| V <sub>prop</sub> +V <sub>prop</sub> | intensification | intensification              | intensification |
| V+V+V                                | continuity      | n.a.                         | continuity      |

(8) *Di Pulu Cocos perumpuan cantik-cantik* (CM)  
on Cocos Islands girls pretty-pretty  
‘On Cocos women are very pretty’

(9) *Saya suka kopi panas-panas*  
I like coffee hot-hot  
‘I like my coffee very hot’

Reduplication is a rather robust feature of the TM of BM and BZM in particular; Malay/Indonesian varieties are rich in productive reduplication strategies, and, among Sinitic languages, Hokkien (and other Min languages) make frequent use of reduplication. This leads to typological congruence and incorporation in the new grammar. The effects of such congruence can be seen in the presence of reduplication in another contact variety of the region, namely Singlish, summarized in [table 7.4](#) (Ansaldo 2004b).

Constituent order in reduplication in BM and BZM is quite variable (Ansaldo and Matthews 1999; Khin Khin Aye 2005): for example, adverbial reduplication can be found both pre- and post-verbally. As shown below, this type of variation is quite frequent in BM and suggests that both the Sinitic and the Malay patterns are possible. Triplication, found in Singlish as well as Hokkien and other Min varieties, is a typologically rare feature that clearly suggests Sinitic influence (Ansaldo 2004b).

Typological congruence between Sinitic and Malayo-Portuguese reduplication patterns also leads to a significant restructuring in Makista, where they almost completely replace the functions of inflectional morphology. For example, in Makista, nominal (full and partial) and adjectival reduplication are used instead of number inflection (Ferreira 1996: 231; Ansaldo and Matthews 2004): *home* ‘man’ > *hóme-hóme* ‘men’; *casa* ‘house’ > *casa-casa* ‘homes’; *nhónha* ‘woman’ > *nho-nhónha* (or *nhónha-nhónha*) ‘women’; *nhum* ‘boy’ > *nhu-nhum* (or *nhum-nhum*) ‘boys, lads’. Nominal reduplication also occurs in Makista with compound nouns: *fula-papaia* ‘papaya flower’ > *fula-papaia-fula-papaia* ‘papaya flowers’; *pê-di-vaca* ‘cow’s foot’ > *pê-di-vaca-pê-di-vaca* ‘cows’ feet’. Adjectival reduplication can indicate plurality on the noun it modifies: *pa tudo nôsso amigo china-china*... [for all our friend Chinese-Chinese],

‘for all our Chinese friends ...’ (Ferreira 1996: 213); *maior parte são gente grande-grande* [greater part are people big-big], ‘it’s mostly important people’ (Tomás 1988: 39). Reduplication patterns also occur in the verbal domain in Makista: adverbs can be reduplicated to modify the verb (10 to 12) and verbs can be reduplicated to indicate intensification and repetition (13 and 14):

- (10) *Ele azinha-azinha subí escada* (Ferreira 1996: 245)  
 he fast-fast go-up stairs  
 ‘He quickly went upstairs’
- (11) *Cedo-cedo sai di casa* (Ferreira 1996: 313)  
 early-early go-out from home  
 ‘(To) leave home early’
- (12) *Iou vagar-vagar já vai riva cozê mizinha* (Ferreira 1996: 48)  
 I slow-slow PFV go top boil medicine  
 ‘I slowly went up to boil some medicine’
- (13) *Venâncio...na cama bulí-bulí su corpo* (Ferreira 1996: 67)  
 Venancio in bed move-move his body  
 ‘Venancio... moved his body about in bed’
- (14) *Testa suado-mulado, coração batê-batê...* (Ferreira 1996: 68)  
 forehead sweat-mottled heart beat-beat  
 ‘His brow mottled with sweat, his heart beating away ...’

Note that reduplication also applies to numerals:

- (15) *Unga-unga ta falá.* (Ferreira 1996: 72)  
 one-one PROG speak  
 ‘(They) are speaking one by one’
- (16) *Unga-unga já virá vai casa* (Ferreira 1996: 86)  
 one-one PFV return go home  
 ‘They returned home one by one’
- (17) *Êle largá pê pulá dôs-dôs degrau* (Ferreira 1996: 72)  
 he left to jump two-two step  
 ‘He left jumping two steps at a time’

In both nominal and verbal domains, many features observed in Makista could be attributed either to Malayo-Portuguese or to Sinitic, in particular Yue and Min varieties (Baxter 1985; Ansaldo and Matthews 2004). For example, nominal reduplication of the type illustrated above has been found in several Asian Creoles and is generally attributed to Malay influence (Holm 1989).<sup>16</sup> The partial reduplication of nouns, where only the first syllable is reduplicated, can be more specifically attributed to Malay influences as this reduplication type is particularly frequent in Malay varieties (Ansaldo and Matthews 2004); indeed partial reduplication of nouns does not appear to be attested in Indo-Portuguese, although there is partial reduplication of certain adjectives (Clements 2003). On the other hand, given that adverbials (reduplicated or not) are post-verbal

in Portuguese and Malay, neither internal development nor Malay influence predicts their pre-verbal position in Makista. Instead, as argued by Ansaldo and Matthews (2004), the pre-verbal adverbial patterns in (10 to 12) above provide the clearest evidence for Sinitic influence in Makista reduplication, since pre-verbal reduplicated adverbs are a characteristic feature of Sinitic grammar: for example, an adjective (or stative verb) can be reduplicated before the verb to serve as an adverbial in Cantonese (Matthews and Yip 1994).

Both Sinitic and Malay support the evolution of numeral reduplication used in a distributive sense; in terms of constituent order, as in the case of the adjective, the pre-verbal position reveals more direct Sinitic influence (Matthews and Yip 1994; Sneddon 1996; Ansaldo and Matthews 2004). One possibility suggested in Ansaldo and Matthews (2004) is that the Malay pattern here hybridized with Chinese syntax to produce the Makista structure; numeral phrases are thereby assimilated to the general Chinese pattern whereby reduplicated adverbials precede the verb.

*iii. Word order in NP* One aspect of the word-order variation just mentioned can be found in the order of nouns and demonstratives. In BM (18a) and BZM (18b), demonstratives tend to precede the noun. This feature, typical of many contact-varieties of Malay, has been attributed to Sinitic influence, considering that Sinitic has [Dem-N] pattern, while Malay follows [N-Dem] pattern:

- (18a) *itu buk*  
       'that book'  
 (18b) *itu hari*  
       'that day'

Variation does occur in the order of these elements: a majority of Lim's informants (1988: 47) reportedly use the Sinitic-like order, with the Malay-like order featuring in a minority of speakers. In conversation these pronouns often appear in a contracted form, such as *ni* and *tu*.

CM follows the same [Dem-N] order; at the same time, it has maintained a more Malay/Indonesian-like tripartite distinction in locative pronouns, with two different distal forms:

- (19a) *sini* 'here'  
 (19b) *sana* 'there'  
 (19c) *situ* 'there out of sight'.

The difference between the two distals can be seen in example (20):

- (20a) *buk di sana*  
       'the book is there', pointing at a book on the shelf  
 (20b) *buk di situ*  
       'the book is (in) there', pointing to a drawer where the book has been placed,  
       not necessarily far but not visible.

Table 7.5. Polyfunctional ‘have’ in BM/BZM, Malay, and Sinitic

| Language/Function | BM (and BZM) | Malay | Sinitic |
|-------------------|--------------|-------|---------|
| Existence         | ✓            | ✓     | ✓       |
| Possession        | ✓            | ✓     | ✓       |
| Aspect/emphasis   | ✓            | ✓     | ✓       |

This tripartite system is typical of Malay, though usually *sana* and *situ* are swapped there (Aone van Engelenhoven, p.c. October 2007).

iv. *Zero copula* A typologically well-attested correlation, namely between zero copula and attributive predication (or property verbs), is found in BM, as well as BZM and CM. There is no overt verb ‘to be’, and no clear difference between verbs and adjectives can be drawn (Lim 1988), e.g. *dia gemok* ‘he fat’ (he is fat); *dia banyak bese* ‘he very big’ (he is very big). The conflation of these traits is found in Sinitic as well as Malay varieties. The strength of Sinitic–Malay congruence in this feature can be better appreciated in Makista grammar which also selects predominantly zero copula structures, resulting in the absence of the verb ‘to be’ which is typical of Portuguese (and other Standard Average European languages), e.g. *este criada bo* (Port. *Esta criada é boa*) ‘this servant (is) good’; *ovos caro!* ‘(the) eggs (are) expensive’ (Arana-Ward 1977: 79). In this sense, the copula is not ‘lost’ in Makista; it is simply not selected because equivalent structures, in which zero copula and property verbs combine, are stronger in the TM.

v. *Existential constructions* The polyfunctional nature of the Malay verb *ada* ‘to have’ has been noted in a number of contact Malay varieties (see Adelaar and Prentice 1996; Khin Khin Aye 2005). Existential constructions for the functions listed in table 7.5 are used both in Malay and Sinitic.

The three functions of Existence, Possession, and Aspect are exemplified below for BM:

- (21) *Ada banyak tempot lagi* (Ansaldo and Matthews 1999: 57)  
have much time more  
‘There is plenty of time’
- (22) *...mana ada anak* (Ansaldo and Matthews 1999: 57)  
how have children  
‘... how can you have children?’
- (23) *Gua ada makan* (Lee 2000: 37)  
I have eat  
‘I did eat’

The first two functions are clearly motivated by polysemy and can be found in both Sinitic and Malay grammars (Ansaldò and Matthews 1999). In its emphatic function, *ada* and its Sinitic counterpart *u* (Hokkien) again show congruence; indeed they can be used on their own as assertive markers, meaning ‘yes’.

In BM narrative, *ada* is used in sentence-initial position in storytelling (Ansaldò and Matthews 1999: 57), a discourse feature found in Singaporean BZM speakers as well (Khin Khin Aye 2005). While this usage is possible in both Malay and Hokkien, it seems far more frequent in the latter, and in Sinitic in general. In addition to the features above, the suppletive negative existential construction *tak* [NP] is consistently used in BM, perhaps a calque of Hokkien *bo* ‘not have’ (Ansaldò and Matthews 1999: 58; see also Bodman 1955), while Malay uses *tak ada* (or *tada*). This parallel is also noted for Singapore BZM (Khin Khin Aye 2005). Here the congruence-induced selection of existential patterns in BM grammar is clear.

*Ada* can also be found as a different aspectual marker indicating continuity and continuous aspect, often in conjunction with sentence-final *lagi*, a typical Malay marker of continuous aspect (Mintz 1994), as shown in (25) and (26). Such co-occurrence with *lagi* suggests Malay influence, even though it is true that the Hokkien existential marker *u* can be used to mark continuous aspect (Bodman 1955; Khin Khin Aye 2005: 270).

- (24) *Dia ada marah* (BM, Lee 2000: 38)  
s/he have angry  
‘S/he is angry’
- (25) *Dia ada makan itu apple lagi* (BM, Lee 2000: 40)  
s/he have eat that apple still  
‘S/he’s still eating the apple’
- (26) *Anak-anak itu lagi lari di sini* (CM)  
children those CONT run LOC here  
‘Those kids are running over here’

Khin Khin Aye (2005) does not report the use of *lagi* in BZM and argues that this indicates stronger Sinitic influence in the evolution of Singapore BZM. It may be the case that Singapore BZM has converged more towards Hokkien than BM, considering the numerical dominance of Sinitic varieties in the ecology of Singapore. However, the analysis presented in Khin Khin Aye (2005) is based on the BZM spoken by people of Chinese background, and it is not surprising therefore to find stronger Sinitic influence in their speech.<sup>17</sup>

*vi. Possessive constructions* Possessive constructions of the type [possessor + *punya* + possessee] are common in BM and Singapore BZM, and constitute a robust PDM feature, as noted in chapter 6. They are typically built on the use



of *punya* 'to possess', realized as *pia* or *mia* in colloquial BM speech. Besides indicating possession (27), in BM and BZM *punya* can be used for two more functions: for attributive predication (28) and as a relative marker (29).

- (27) *Dia punya bilik* (BM)  
 s/he POSS room  
 'his/her room'
- (28) *Betol punya hap* (BM)  
 good POSS match  
 'a good match'
- (29) *Gua pukol punya itu orang* (BM, Lim 1988: 50)  
 I hit POSS that man  
 'the man I hit'

This type of possessive construction based on the possessive marker *e*, extended in function to a marker of modifying phrases, is found in Hokkien and other Min languages (Pakir 1986; Lim 1988; Ansaldo and Matthews 1999: 56), as shown in (30) and (31).

- (30) *Khaq tua e cabo*  
 more old POSS woman  
 'an older woman'
- (31) *Gun ce e chia* (Bodman 1955: 238 in Khin Khin Aye 2005: 263)  
 we sit POSS car  
 'The car we sit in'

As noted in Ansaldo and Matthews (1999), the syntactic pattern most recurrent in BM follows that of Sinitic [Possessor + Possessee] and not Malay [Possessee + Possessor]. Variation in this word order may occur: Lim (1988: 47) reports high variation in constituent order of *punya* constructions when used to mark possession and attributive predication, indicating that both the Sinitic and the Malay patterns are used in BM, with the former being the dominant one. The BM possessive marker thus most closely resembles that of Hokkien *e* in its functional range, as also noted for Singapore BZM in Khin Khin Aye (2005). In particular in its attributive and relational usage, it appears that the semantic functions of the Hokkien possessive marker *e* have been selected in the restructuring.

The similarities above have often been contrasted to Malay varieties where possession is typically marked by juxtaposition (Lim 1988; Khin Khin Aye 2005). Indeed, while in CM, *punya* constructions are possible, they are not as common and have a more limited range of functions. The use of possessive *punya* 'to have/to possess' is often optional in CM (34) since unmarked juxtaposition of the Malay type occurs. In fact, *punya* is not attested in elders' speech,<sup>18</sup> while *ada* is more commonly used to express possession (24 to 26).

- (32) *Saya punya bapak sobat* (CM)  
 I poss dad friend  
 'My dad's friend'
- (33) *Jukung dia di situ* (CM)  
 boat s/he LOC here  
 'His/her boat is here'
- (34) *Matt banya bagus* (CM)  
 Matt very good  
 'Matt is very nice'

Equally common in CM is the use of the suffix *-nya* to mark possession: *ini dokternya* 'this is the doctor's; *namanya Lisa* 'her name is Lisa'. Note that *-nya* can also mark specificity in CM: *ini bukunya* 'this is *the* book (I was talking about)'. The presence of this suffix is also reported in Ambonese Malay and other contact Malay varieties (Adelaar 1996b: 183 fn.5; Uri Tadmor p.c. 2002). The suffix is also very productive and versatile in more standard Malay/Indonesian varieties (Mintz 1994).

The possessive construction based on *punya* shows syntactic and semantic parallels in BM and BZM, which can reasonably be attributed to transfer of Hokkien features into BM and Singapore BZM, and thus represents a clear Sinitic influence in PDM varieties. Its use is also found in CM, though it is optional and restricted to the core function of marking possession. For relative clauses, CM uses the Malay relative marker *yang*.<sup>19</sup>

vii. *Causative constructions* Causative constructions in BM are based on two Malay verbs: *kasi* 'to give' and *buat* 'to make', which occur in pre-verbal position. The lexical verb *kasi* 'to give' is somewhat more frequent and is found in three functional contexts in BM (and BZM): (a) Benefactive (35), (b) Causative (36), and (c) Passive or Agent-marker (37):

- (35) *Dia beli itu baju kasi gua* (BM, Lim 1988: 20)  
 s/he buy that dress give I  
 'S/he buys that dress for me'
- (36) *Dia pekek-pekek kasi gua tapranjat* (BM, Lim 1988: 21)  
 she scream give I startle  
 'Her scream startled me'
- (37) *Dia kasi gua pukul* (BM, Lim 1988: 21)  
 he give I hit  
 'He was hit by me' or 'He caused me to hit him'

BZM functions are identical to BM. The semantic parallel between *kasi* and Hokkien *ho* 'to give' is rather clear, as described in Lim (1988: 20), Ansaldo and Matthews (1999: 53), and Khin Khin Aye (2005) for Singapore BZM. *Ho* is often used in serial verb constructions as a Benefactive (or Dative) marker

(Bodman 1955; Yap and Iwasaki 2003); in addition, it is used as a Causative and Agent marker in Hokkien, as in (38) and (39).

- (38) *Li ho gua khua* (Bodman 1955: 310)

you give me see

'You make me see/realize'

- (39) *I ca-zit ho laj phaq*

he yesterday give people beat

'He was/got beaten by someone yesterday'

In CM, *kasi* is predominantly used to mark Causative/Permissive. It is at variance with *bikin*, of more generic Malay/Indonesian origin (see (41) and Adelaar 1996b: 178).

- (40) *Guru kasi anak itu pulang rumah* (CM)

teacher give child that return home

'The teacher let the child go home'

- (41) *Makanan semalam bikin saya sakit*

food yesterday do I sick

'Yesterday's food made me sick'

One has to bear in mind, though, that Benefactive/Causative *kasi* is ubiquitous among contact and vernacular Malay varieties (Pakir 1986: 176–7); in this sense it is then really the passive function that should be closely attributed to Hokkien influence. Interestingly, this function is not found in CM, where passive is marked by *kena*. In fact, even in BM and Singapore BZM, it is *kena* that marks passive (see viii below), while *kasi* as passive marker is extremely infrequent. Therefore, the clearest possible indication of a Hokkien influence in *kasi*-constructions also pertains to the weakest semantic feature, which suggests that congruence, rather than exclusive influence from one particular adstrate, is involved in the selection of causative constructions. Finally, note that in all these varieties, *kasi* can also function as a full lexical verb (42).

- (42) *Kasi saya itu rokok* (CM)

give I that cigarette

'Give me that/the cigarette(s)'

viii. *Kena passive* A feature that has attracted quite a bit of attention in studies of contact Malay varieties is the use of *kena*, a Malay word with a connotation of abstract contact (e.g. 'to incur', Lim 1988: 21), as a marker of non-volition and/or passive constructions (43 to 45).<sup>20</sup> In addition, its semantics implies a negative, unavoidable encounter.

- (43) *Nkoh gua kena orang tipu* (BM)

brother I kena people cheat

'My brother was cheated'

- (44) *Ini budak kena pukul* (BZM, Khin Khin Aye 2005: 279)  
 This child *kena* beat  
 'This child was beaten'
- (45) *Saya kena pintu.* (CM)  
 I *kena* door  
 'I was hit by a door.' or 'I (accidentally) hit a door'
- (46) *Saya kena kerja.* (CM)  
 I *kena* work  
 'I was made to work' or 'I had to work (but I didn't want to)'

As is clear from the above CM examples, a passive reading can be ascribed to these sentences, but we could also read them in an active way. When *kena* is assigned an abstract interpretation, viz., 'to suffer something', then (43) and (44) can be interpreted, respectively, as 'my brother ran into people who cheated (him)' and 'the child incurred a beating'. Distinguishing clearly between non-volition and obligation is a tricky issue: obligation could indeed be seen as one interpretation of non-volition, and one must clearly distinguish between a passive reading (47 and 48) and a passive construction (49). In a sentence like (46), we can only disambiguate through context, i.e. interpretation, but we cannot really know from the sentence alone whether the person was made to work or had to work because he was broke. In this sense it seems that the Malay properties of *kena* account for its usage in BM, BZM, and CM equally well and that convergence with Hokkien may have been exaggerated (see Pakir 1986; Ansaldo and Matthews 1999; Khin Khin Aye 2005). Note that *kena* is used in colloquial Malay and Indonesian to mark non-volition and obligation, e.g. *dia kena pukul* 's/he got hit'. A parallel structure is found in Hokkien (Lim 1988: 21): *i tioq phak*. *Tioq* also indicates non-volition as well as obligation. Obligation and/or necessity can be expressed through *kena* constructions in several varieties of Malay, as exemplified for CM:

- (47) *Saya kena masak*  
 I need/must cook  
 'It's my turn to cook' or 'I need to cook'
- (48) *Paman kena mancin*  
 I need/must fish  
 'I need to go fishing'

Lim (1988) and Khin Khin Aye for BZM (2005) suggest that the semantic similarity and functional proximity between BM, Malay, and Hokkien illustrate substrate reinforcement. Their position amounts to what is interpreted here as typological congruence.

*ix. Isolating morphology* BM has hardly any productive morphology of the Malay/Indonesian type: Pakir (1986: 132) notes that applicative suffixes such as *-kan* are the exception rather than the rule in BM, and the texts reported in Lim

(1988) show almost total absence of productive morphology. Recall that we are not comparing BM with Standard Malay. Such a comparison would not be realistic, since BM evolved prior to the standardization of Malay, aside from the fact that literary genres of Malay do not feature prominently in the ecology. Taking this into consideration, we should not expect the full array of morphological devices associated with Malay to enter the TM in which BM evolves. For the present, it is noteworthy that, while BM and BZM are clearly typologically isolating, a feature that has been attributed to convergence with Hokkien isolating typology (Lim 1988; Khin Khin Aye 2005), CM offers an example of a contact Malay variety in which affixes typical of Malay/Indonesian languages are found (Adelaar 1996a).

In relation to the ‘passive’ construction discussed above, CM has the typical Malay passive prefix *di-* (49), while *kena* marks non-volition or obligation (47 and 48). The agent may or may not be present in these constructions.

- (49) *Saya di-potong (Lisa)*  
 I PASS-cut (Lisa)  
 ‘I was cut (by Lisa)’

A similar function is found in *ter-*, a prefix that detransitivizes the verb, e.g.

- (50) *Saya ter-luka*  
 I ter-hurt  
 ‘I was (accidentally) hurt’

Sentences of this kind usually imply something accidental: the Topic (*saya*) is non-volitional or out of the subject’s control (see also Adelaar 1996b: 173). Though affixes are often at variance with periphrastic constructions in CM, it is a fact that morphologically CM has a less isolating typology than BM and Singapore BZM. This may reflect particular influence from Indonesian language varieties. However, as shown in table 7.6, some processes of affixation found in CM are not really productive and might better be treated as lexicalized forms (Ansaldò 2002).

As is evident from table 7.6, despite the limited productivity of some of the affixes, CM is morphologically much richer than BM. Obviously, if we assume that the female component of the BM community spoke varieties akin to CM, a shift to isolating typology must have occurred in the history of the language. Crucial in explaining this may be the fact that, of the morphological material found in CM, very little is obligatory, and much is unproductive; in other words, frequency is low. As in the case of the Accusative in SLM (chapter 6), this can result in lower selection rates in the restructuring process. Moreover, as indicated in chapter 3, CM evolves predominantly with a Malay/Indonesian TM, without the presence of strongly isolating languages such as Sinitic, which explains why some morphological processes are still competitive (i.e. productive). This suggests that, while Sinitic is clearly relevant in the evolution of some contact Malay varieties, other evolutionary paths also took place.

Table 7.6. *Morphological material in Cocos Malay*

| Affix type                                                  | Example                                        | Notes                                             |
|-------------------------------------------------------------|------------------------------------------------|---------------------------------------------------|
| <i>V-an</i> > N                                             | <i>main</i> ‘to play’ > <i>mainan</i> ‘games’  | quite productive                                  |
| <i>ke-V-an</i> > N                                          | <i>kemainan</i> ‘toys/games’                   | less productive                                   |
| <i>per-V-an</i> > N <sub>abstract</sub>                     | <i>atur</i> ‘to rule’ > <i>peraturan</i> ‘law’ | less productive                                   |
| <i>V<sub>intr</sub>-kan</i> > <i>V<sub>trans</sub></i>      | <i>belikan</i> ‘to buy for someone’            | productive                                        |
| <i>meng-V-kan</i> > <i>V<sub>trans+</sub></i> <sup>21</sup> | <i>mentakut(kan)</i> ‘to fear’                 | often <i>meng-</i> and <i>-kan</i> occur together |
| <i>ber-V</i> > <i>V<sub>trans+</sub></i>                    | <i>berbicara</i> ‘to talk about something’     | unclear                                           |
| <i>baku-V</i> > <i>V<sub>reciprocal</sub></i>               | <i>bakubeli</i> ‘buy for each other’           | productive                                        |
| <i>di-V</i> > passive                                       | <i>dipotong</i> ‘to be cut’                    | productive                                        |
| <i>ter-V</i> > non-volitional                               | <i>terluka</i> ‘to be hurt’                    | less productive                                   |

### 7.3.4 Summary

The features presented in this section are by no means exhaustive of the grammars of BM and Makista, but they sketch a sufficiently accurate picture of salient typological traits. In it, I have highlighted the significance of congruence against privileging particular adstrate languages over others in traditional analyses of CLF. For example, Ansaldo and Matthews (1999), in highlighting some aspects of BM grammar, pay closer attention to the Sinitic rather than the Malay aspect, and the result of such an analysis appears to suggest stronger influence of Hokkien for the specific properties. Likewise, Khin Khin Aye (2005) stresses the Hokkien parallels in her relexification-based take on BZM, but her analysis relies heavily on Chinese speakers of BZM in the first place, and considers mostly standard Malay features as a point of comparison, which, as has already been pointed out, is methodologically inappropriate. It is of course understandable that in these analyses, the comparison to Malay varieties appears somewhat remote, as it is difficult to establish with certainty which variety of Malay or Indonesian should be used. As already noted in several places above, there were most likely several Malay/Indonesian vernaculars as well as contact Malay varieties present in the matrix. In such contexts, we must therefore refrain from drawing hasty conclusions about the exact origins of specific features, and acknowledge instead the pluralistic composition of the matrix. As already suggested in Pakir (1986), the Hokkien element in BM should not be overstated since much of its grammar is also Malay.

As shown in this section, a TM-based approach reveals that the strongest pattern in the evolution of BM is in fact one of congruence: of the grammatical features discussed in section 7.3, the following are all a clear result of congruence between Hokkien and vernacular Malay/Indonesian typology:

Topic–Comment structure, reduplication, zero copula, existentials, causatives, and passives. The only feature that is clearly more Sinitic than Malay in the grammar is the possessive construction. It is possible that the discourse function of the existential marker is also slightly more Sinitic, while for causatives and passives the evidence so far is not conclusive. Word-order variation reveals simply that both adstrates contribute to word order patterns, though the Sinitic order is somewhat more frequent. This might vary depending on the nature of the data collected, as shown in the case of the Singapore BZM spoken by ethnically Chinese Singaporeans described in Khin Khin Aye (2005). Typologically, we see that BM is far more isolating than CM, as the latter has less Sinitic and more Malay/Indonesian elements in its grammar. This has been related to the fact that, even within PDM varieties, we may want to allow for different patterns of evolution (Paauw 2003).

The lexicon of BM is predominantly Malay with a few Sinitic items, which relate to kinship, family, and terms of address, and show maintenance of Chinese culture in the home, while the elaborate set of terms of address is a general feature of Malay/Indonesian languages. Phonologically BM is basically Malay, a fact that led Pakir (1986) to claim that the Sinitic influence might have entered the language at a later stage. This is, however, unlikely in the light of the founder principle: specific Sinitic elements such as possessive constructions in BM as well as other PDM varieties are quite likely to be very archaic features (see chapter 6). What really matters here is that an important factor in the formation of BM is congruence. In almost all instances where typological similarities are found, the relevant features have been selected into the emergent variety. Where there is no congruence, selection can go either way. For example, the core of Sinitic words may indicate a patriarchal community where males dominated; at the same time, the phonological congruence to Malay indicates that strong accommodation towards local languages of Malaysia and Indonesia (and perhaps active differentiation from non-localized Chinese traders) was important in the evolution of BM. In this sense, the congruence that occurs between the linguistic systems can be seen as a manifestation of the social accommodation that occurs between the speakers.

The argument that if BM were simply a blend of Hokkien and Malay features it would show more Hokkien features than merely some “structural hiccups” (McWhorter 2007: 231–3) rests on a basic misunderstanding of elementary mathematics. If every language has one grammar, for the sake of simplicity ‘1’, then, in the TM of BM, we have 1 for a BZM variety and 1 for Hokkien. Note that, in reality, the picture is not as simple, as there were several Sinitic varieties and Malay/Indonesian languages involved in the ecology of BM. So we have something like:  $1 + 1 + 1 + 1$  (impressionistically) = 4. Now, if we agree that all human grammars are roughly the same in nature, then any single language’s grammar has the value 1; therefore it cannot have grammar 4 (or 3

or 2), because any language only has grammar value 1. This is why BM does not have all the Hokkien features McWhorter desires.

The additional comparison with Makista illustrates how congruence can explain what appear to be (a) innovative features and (b) simplification processes. Reduplication in Makista has evolved because it is a frequent typological feature of Malay and Sinitic alike, which were both involved in the TM. In Malay and Sinitic grammars, in particular Min, reduplication strategies occur in several domains and with numerous functions; this means that type-frequency within each grammar is high. Moreover, reduplication patterns in Malay and Hokkien are to a large extent comparable, as can be seen in the iterative and intensifying functions of verbal reduplication; this means that type-frequency is further reinforced in the matrix. Instead of claiming that Makista has ‘lost’ number inflection, from the point of view of the TM we can say that nominal and adjectival reduplication take over the function of plural marking and render inflectional morphology irrelevant. Likewise, zero copula (or equative constructions) evolved because of type-frequency in the TM, as it is a solid feature of Malay and Sinitic, and renders the copula verb redundant. We could of course think of this as structural reduction, in the sense that there is one less overt item in the grammar; but we could also think of this as a semantic addition, in the sense that in equative constructions adjectives also behave verbally. As already argued elsewhere (Ansaldo and Matthews 2001), simplicity and complexity are not really objectively quantifiable and should be regarded as language-specific notions. The fact that there is less structural hybridity in Makista than in BM (see Ansaldo and Matthews 2004) can be attributed to the conflation of two factors: (1) the possibility of congruence in the TM of Makista is reduced, because there is more distance between the typologies involved; and (2) social isolation of the community is stronger, due to political and historical factors (see section 3.4). In Makista we thus find more retention of Malayo-Portuguese features, the ancestral variety of the community.

## 7.4 Conclusions

It appears that in all the communities described above, multilingualism and multiculturalism were defining features of social positioning and economic advantage. In all three cases, their position as intermediaries meant sufficient adaptation to the environment in order to benefit from it, while at the same time it demarcated them as a discrete cultural entity. The differentiation was a socio-economic one in the case of the Babas, while it was a predominantly political one in the case of the Macanese vis-à-vis China and Portugal. In the case of the Malays of Sri Lanka, the ‘separation’ was predominantly social, through endogamy. As suggested above, linguistic differentiation is a very efficient identity marker. However, it does not come alone, as it is typically



part of cultural adaptation and innovation. Finally, what is striking in all cases is the fact that identity alignments are continuous, recurrent patterns in contact ecologies; they are expressions of multilingual identity and multicultural adaptation triggered by the need of a group to reposition itself within evolving socio-historical contexts.

In this chapter I have proposed that the evolution of a new variety is the result of a process of identity alignment in a community. Identity alignment involves innovative construction of a new grammar based on multilingual resources, and is part of a more general process of new cultural identification that involves intra-group focussing and inter-group differentiation. On the other hand, differentiation does not mean total segregation, because the multilingual and multicultural resources of the group always allow contact to be maintained. Through this particular social positioning, the group often achieves an advantageous position in the political ecology. Identity alignment thus suggests that CLF is not the result of some exceptional situation in which speakers were denied the social, cultural, or cognitive settings necessary for a 'normal' acquisition of a target; rather, it is a creative process in which speakers have agency in the transmission and transfer of linguistic features. Processes of selection and replication from a diverse pool underlie both CLF and identity alignment; frequency plays an important role in the selection process and, where it does not prevail, founder effects can be observed, as in many Asian-Portuguese Creoles that are characterized by substantial maintenance, rather than innovation. Admixture is thus a perfectly normal outcome of CLF, as the combination of acts of maintenance and divergence that defines an emerging community in a new ecology. Admixture is the outcome of multilingual practices in linguistically diverse environments; CLF is the result of cultural focussing through identity alignment.

## Notes

- 1 I will use *Makista* when referring to the language and *Macanese* for the community and their culture.
- 2 These findings come from a small-scale survey of the Colombo Malays conducted in 2006 (Lisa Lim p.c. July 2006).
- 3 Indeed, their language, Baba Malay, has been adopted by other groups such as the Chitty Malays, a group of mixed Indian-Malay origin (Alan Koh p.c. 20 Jan 2008).
- 4 This was not necessarily the case for the ships leaving from Portugal, but the Asian routes were very dependent on non-European sailors (Hugo Cardoso p.c. 18 Nov 2008).
- 5 Though the Manila Trade between Macau, China, and the Spanish in the Philippines was extremely lucrative to the Macanese, their relationship to the Spaniards was one of mistrust. The Spanish tried to use their influence in the Philippines to establish direct links with China but Macanese authorities refused any interference in their monopoly (Gunn 1996: 23).

- 6 The largest Hokkien community outside China is still found in Bangkok today.
- 7 These findings also come from the small-scale survey of the Colombo Malays conducted in 2006 (Lisa Lim p.c. July 2006).
- 8 Canarim refers to the Kannara coast of India, including Karnataka, Goa, and parts of Maharashtra (Hugo Cardoso p.c. 2007).
- 9 Other Sinitic varieties were also present, in particular Teochew, another Min language, which bears structural resemblance to Hokkien. However, this variety seems to have left little trace in the ecology of Singapore (or is perhaps under-investigated); for example, though the Teochews were demographically dominant, the Hokkiens seem to have had a stronger linguistic input, possibly a reflection of their higher prestige in the community (Lim 2007).
- 10 Recall that, in this framework, it is actually occurring features that count, not potentially existing ones; in other words, it is usage, not competence, that defines the content of the TM.
- 11 This was indeed the case in Batavia, which may also explain why the Dutch language never took hold in the region. Portuguese was not only spoken in some form by many servants and slaves, but was also familiar to the Dutch merchants from the salt trade between Amsterdam and Portugal (Taylor 1983).
- 12 The representation of Malay data follows conventional Malay scholarship. Hokkien and Makista data are represented according to the source. Cantonese data are given in the Yale romanization system. Hokkien tones are omitted as they are not relevant for this study.
- 13 While Hokkien lexical influence is to be found in culturally salient domains and thus indicates a strong presence of Chinese values, other influences are also found. Pakir (1986: 93) suggests that, numerically, Hokkien borrowings do not exceed those of Arabic origin (very high in Malay), e.g. *akal* 'ingenuity', *adat* 'custom'; moreover, there are also Sanskrit and Tamil loans that BM inherited through Malay varieties.
- 14 For an analysis of the CM lexicon, as well as aspects of grammar, see also Adelaar (1996b).
- 15 Despite a number of slaves of African origin, hardly any trace is found of African vocabulary, which may indicate that black slaves in Macau spoke a form of Portuguese (Amaro 1994: 39).
- 16 This conclusion is probably correct, although the possibility of influence by other Austronesian languages cannot be ruled out. Interestingly, Clements (2003) shows that some forms of nominal reduplication exist in Gulf of Guinea varieties (São Tomense, Fa d'Ambu) as well as in Indo-Portuguese varieties (Korlai, Daman). This raises the possibility that nominal reduplication was already present in vernacular varieties that the Portuguese may have used in India, Melaka, and Macau. Indeed, Hugo Cardoso (p.c. 18 Nov 2008) notes that, according to the corpus available to Schuchardt (1883), Diu Indo-Portuguese made widespread use of nominal reduplication as a pluralizing strategy.
- 17 In this sense Khin Khin Aye (2005) overstates the case for Hokkien in her relexification approach; this is clear if we consider that the comparison between BZM and standard Malay is questionable in terms of substrate analysis. Standard Malay is a recent creation and differs in a number of features from colloquial Malay. Moreover, the diversity of Malay/Indonesian varieties spoken in the ecology of BZM is lost in the standard variety. Existential constructions of this kind are found in both Sinitic and Malay varieties, and there is clearly typological congruence in BM and BZM

which is not always easy, nor necessary in an evolutionary account, to trace back to one specific origin.

- 18 Though it is difficult to establish whether it should be considered a generational or an age-graded feature.
- 19 *Dia lapas yang kecil/Yang kecil dia lapas* 'the small ones he lets go'.
- 20 The term *kena* is not fully grammaticalized and occurs in its verbal usage in most varieties of Malay (Mintz 1994; Khin Khin Aye 2005).
- 21 The use of '+' indicates that the affix increases transitivity, whether of an attributive verb, an intransitive, or a transitive verb.

## 8 Pidgin ecologies of the China coast

---

*Englishman no can talkee Chinee, he no plenty smart inside. Allo Chinaman talkee Englishee all plopa – all-same my*

Charles Leland (1892: 111)

In this chapter we turn our attention to Southern China, where speakers of Asian-Portuguese, English (and other Western) vernaculars, and Sinitic varieties came into contact for the purpose of trade. Because of the extremely protectionist nature of the Chinese empire, and due to the fierce competition between Portuguese, Dutch, and British trading companies, the contacts between the Eastern and Western parties were in fact very limited in nature and confined to few social settings, as already seen in the history of Macau (section 3.4 and chapter 7). Moreover, most of the official contacts were regulated by a small group of intermediaries. In such confined contexts, limited contacts centred around very specific interests such as commerce often led to the formation of so-called ‘pidgin’ languages. These are restricted codes that serve the purpose of basic communication in a specialized field, as in the case of *Russenorsk*, a Norwegian–Russian pidgin spoken in the periodic settlements of the northern Norwegian coast between Russian and Norwegian whalers. In Southern China, a unique contact ecology evolved around the Pearl River Delta between the city-ports of Macau and Canton<sup>1</sup> (now Guangzhou, see map 8.1). For a long time this was the only area officially accessible to Western ships and it constituted the nexus of the trade between Southeast Asia and Japan as well as the centre of the commerce between Europe and the Far East. In this area, a linguistic environment subject to laws and influences different from anywhere else in the region, a contact language evolved referred to as China Coast Pidgin (CCP) or Chinese Pidgin English. CCP has been called the “mother of all pidgins” for a number of reasons (Li, Matthews and Smith 2005: 17). In particular, it seems to have given the very name of *pidgin* to functionally restricted codes that arise in prevalently bilingual communities, through a Chinese-influenced pronunciation of the word *business* (Reinecke 1937). CCP has also been associated with the spread of English pidgins in the Pacific, and it has been considered an example of a ‘stable’ linguistic system



Map 8.1 Pearl River Delta

(Tryon, Mühlhäusler and Baker 1996). Moreover, the fact that it was written down and actively studied (section 8.2.4) renders it rather unique in the history of pidgin development.

In this chapter, I describe the evolution of this variety in relation to theories of language formation, and focus on the nature of microclimates within an evolutionary theory of contact language formation. The fundamental issues here relate to (a) the type of social interactions that lead to the evolution of pidgins, and (b) the role of the pidgin in the wider ecology of language evolution. The first point is relevant to understanding how we explain the grammatical aspects of a pidgin language. It is common knowledge that pidgins look ‘simple’

compared to the languages spoken natively by their users. How does such simplicity evolve? Is simplification a necessary component of second-language acquisition? The second point relates to an even better known generalization, namely that pidgins are the origins of Creole languages. Do the facts leading to the emergence of CCP support the now-disputed pidgin-creole evolutionary hypothesis, or is it yet another nail in the coffin?

## 8.1      Socio-historical background of Europe–China relations

In order to appreciate the history of East–West encounters, it is important to realize that Western arrivals in Southern China were not seen as intrepid adventurers, lucrative business-partners, or bearers of higher knowledge. China's attitudes to non-Chinese, in particular towards populations beyond its Western borders, have always been clear: they were outsiders and barbarians, a potentially polluting element for the superior Chinese culture; they had to be kept at bay at all costs. As explained in [chapter 2](#), Chinese authorities had been intent on limiting and controlling interactions between Chinese and outsiders since the fifteenth century. The same treatment was applied to the Portuguese, the Dutch, and especially the British, as revealed in historical accounts of East–West relations between the seventeenth and eighteenth centuries.

In 1739, Lord Macartney arrived at the Qing court on a 66-gun man-of-war: he was accompanied by one hundred scientists and artists, and brought two vessels loaded with British manufacturing technology. His mission was to request an expansion of trade under published tariffs. The Emperor accepted the gift as tributes from what was, in the eyes of imperial China, a minuscule kingdom of the 'Little Western Ocean' (the Atlantic), politely declined the offer to trade, and said, "I have already taken note of your respectful spirit of submission ... I do not forget the lonely remoteness of your island, cut off from the world by intervening wastes of sea" (Fairbank 1953: 59). Years later, a similar approach was turned down again, this time much more rudely (Fairbank and Goldman 1998: 196–7; Gelber 2007: 64).

Chinese trade did not need the West because it was rich and successful at this point in time, profiting from the overland silk trade, the spice trade with Maluku, and trade with Japan; in this respect, the wool and cotton garments that the British had to offer were pathetic in the eyes of the Chinese. Indeed, Western missionaries and adventurers since the fifteenth century had described China as a rich and impressive country, where material wealth, architectural wonders, and advanced scientific knowledge were to be found (Boxer 1953). The feeling of superiority that the Chinese displayed towards Europeans was well captured by Thomas Taylor Meadows, a British interpreter and student of China, who in 1852 wrote:

The Chinese do habitually call and consider Europeans ‘barbarians’, meaning by that term ‘peoples in a rude, uncivilized state, morally and intellectually uncultivated’ ... Those Chinese who have had direct opportunities of learning something of our customs and culture – they may amount, taking all Five Ports, to some five or six thousand out of the three hundred and sixty millions – mostly consider us beneath their nation in moral and intellectual cultivation. As to those who have had no such opportunities, I do not recollect conversing with one ... They are always surprised, not to say astonished, to learn that we have surnames, and understand the family distinctions of father, brother, wife, sister, etc.; in short, that we live otherwise than as a herd of cattle.

It comes therefore as no surprise that both Portuguese and, later, British traders were kept at a safe distance, isolated in specially defined zones. The general reaction of the Chinese toward the Portuguese in the Pearl River Delta was of simple tolerance: they represented just another band of pirates and traders, a far lesser worry to the Chinese than the numerous barbarian tribes in the north. As explained in [chapter 3](#), the Portuguese were confined to Macau where, officially, they were not allowed to co-habit with Chinese, and were only allowed to trade, periodically, under supervision of Chinese authorities. The British, arriving in the region much later, were, after reaching the Pearl River Delta, first of all ignored, and then sent to Macau, in the hope that the Portuguese would deal with them and, in so doing, spare the Chinese from direct encounters. After a British vessel gunned its way up the Pearl River Delta, British and other foreign merchants were confined to one area of Canton in specially designated buildings, called ‘factories’ (see [section 8.2.1](#)) only for the period of trade.

It was much later, only as a consequence of the Opium Wars (1842–54), that China had to open up several trading ports to allow Western merchants to trade (Fairbank 1953; Fairbank and Goldman 1998). The ‘gun-diplomacy’ of the British reinforced the Chinese view of the ‘redhaired barbarians’, a term first used for central Asian tribes and then extended to Dutch and British traders. After all, the Chinese considered themselves a relatively peaceful society, having entertained mostly diplomatic relations within Monsoon Asia for over five centuries. In his historical analysis of East–West contacts in the nineteenth century, Fairbank (1953) is clear on the fact that, from the Chinese point of view, Westerners were culturally inferior, especially because they were generally traders, a lowly occupation in their society:<sup>2</sup> “The decadent, part half-caste community of Macau remained walled off on its peninsula; the thirteen Factories at Canton were outside the city walls, from which foreigners [Westerners] were excluded.” (Fairbank 1953: 13)

The only group that had managed to break through the iron curtain that the Chinese had put up against foreigners throughout their history were missionaries, in particular the Jesuits, but even they were eventually dismissed. This occurred after they lost the ‘controversy over rites’, a public debate

between scholars of different religious traditions held by the imperial court in 1725, and they became virtually powerless since the rejection by the ‘Son of Heaven’ of the Pope’s ecumenical claims (Fairbank 1953: 14).

## 8.2 The ecology of Macau and the Pearl River Delta

The history of Macau and the Pearl River Delta constitutes probably the most fundamental chapter in the development of East–West contacts in China and the linguistic ecologies that resulted from these encounters. From the point of view of contact language formation (CLF), there are in particular two interesting ‘objects’ of enquiry: (1) Makista, already discussed in [chapters 3 and 7](#), the Asian-Portuguese vernacular with Sinitic elements that became the idiom of the mixed Portuguese–Asian community of Macau; and (2) China Coast Pidgin (CCP), the trade jargon that developed around the Canton area during British colonial rule and later spread to other ports of China and, according to some creolists, into the Pacific. In this section, I address the ecology and evolution of CCP, focussing in [section 8.4](#) on its relation to Makista.

### 8.2.1 *The British in Southern China*<sup>3</sup>

The British arrived in Southern China over a century after the Portuguese. At the time of their arrival (1644–84), several problems were occupying the newly installed Qing Dynasty. Besides the usual border wars against Altaic tribes in the northwest, Chinese authorities were busy dealing with the uprising of the Taiwanese under the ‘pirate’ Koxinga, which, among other things, contributed to the ban on maritime commerce. Chinese authorities were thus unconcerned with the new traders, and happy to use the Macanese as mediators in the initial contacts. Macanese authorities were interested in protecting their monopoly on trade with China and so attempted to confine and control the activities of the British as much as possible. The British, however, resented Portuguese mediation and tried hard to trade directly with Chinese authorities. In 1685, an imperial decree determined that Canton would be opened for trade, and in 1699 the British were allowed to establish a ‘factory’ there (Gunn 1996: 27) – the term is derived from Portuguese *feitor*, a royal agent dedicated to the trade in goods under royal monopoly, a central figure in Portuguese mercantile history since the early fifteenth century (Paviot 2005: 24).

It is only from the latter part of the seventeenth century that the British presence started growing in Southern China, as a result of (a) the refusal by Japan to trade with the British, which led to the latter’s growing interest in the China trade, (b) the recognition by other Western trading powers of the East India Company, and (c) the resolution of the conflict between China and Taiwan, which re-opened the maritime routes to China (Martino 2003).



In terms of linguistic ecology, we should not automatically assume that the arrival of the British meant an influx of English into the region. From the comments of Peter Mundy (Temple 1919), we learn about the difficulties the British encountered in trying to communicate in English with the locals. We also know that the use of a mixture of English and Portuguese in the interactions with local Chinese was common until well into the middle of the eighteenth century (Noble 1762; Hickey 1769 in Quennell 1975). It is obvious that Macau and Portuguese vernaculars played important roles in the early trade between Western merchants and China for a number of reasons. Firstly, a form of vernacular Portuguese had become the *lingua franca* of the coastal regions, due to the predominance of Portuguese ships in the China trade (Martino 2003). It was used amongst Western traders, and was also the best means of communication with Chinese parties until the late seventeenth century (Reinecke 1937: 774–5). To be sure, even the British used “negroes fugitive of the Portuguese” as mediators (The *Chinese Repository*, Anon. 1833, cited in Martino 2003: 21). Moreover, the Portuguese were regarded as the best source of information on Chinese (as well as Japanese) culture and manners, because it was in Macau that the only schools for the study of Chinese language and culture could be found, especially linked to missionary activity (Tamburello 1983). Secondly, among the *casados*, the sons of the Macanese, a class of interpreters grew over time who specialized in mediating between the Chinese authorities and Portuguese officials. They were fluent in written and spoken Chinese, most likely both Cantonese and Mandarin, as they were dealing with local authorities, who would require the former, as well as court officials from the capital, for whom only the latter would be acceptable (Boxer 1973).<sup>4</sup> This linguistic expertise positioned them at the centre of the East–West trade as interpreters and mediators, a role that they held on to until even after the establishment of Dutch and British trading houses led to a decline of Portuguese economic power. Finally, even after the opening of Canton, Macau remained a periodic residence of Western traders, in particular each summer and autumn, when the city would fill up with traders waiting for their boats to be measured or papers to be issued (Gunn 1996: 28).

For these reasons, the role of the English language would have been irrelevant in the first phase of British–Chinese relations (Martino 2003: 24): not only was Portuguese mediation central in this phase but British presence in the region was also severely limited and lacked constant presence, as the British were still very focussed on India and North America until the end of the century. Though it is impossible to rule out English–Chinese linguistic contacts in this early phase, it is much more likely that CCP evolved in a later period, following the ‘Flint case’ and the development of the ‘Canton System’.

### 8.2.2 *The Canton System*

Already in the early days of the Canton Trade, commerce was regulated by Chinese authorities and was confined to a limited number of ports where Western merchants carried out business through local Chinese merchants, among whom corruption abounded. In an attempt to liberalize and improve commerce, James Flint, who resided in China between 1736 and 1762, travelled to Tianjin to address the Emperor about the question of corruption of the local Cantonese authorities (Soothill 1925; Martino 2003: 32). He obtained help in translating a petition on corruption in the Canton Trade and presented it to the Emperor. The consequences of this interference were rather drastic: Flint was imprisoned in Macau for three years and then expelled from China; all ports but Canton were closed to foreigners; the translator of the petition was beheaded; and it was forbidden to teach foreigners the Chinese language, the penalty being the death sentence. This also led to the establishment of the 'Canton System' and resulted in the fact that, until 1842, virtually all power was in Chinese hands, with the movements of Westerners heavily restricted, contact with the local population prohibited except through officially appointed mediators, and an official position of inferiority accepted by Western traders for the sake of economic profit. This is significant because, as a consequence of these measures, the British never gained knowledge of the Chinese language, nor did they set up institutions to study Chinese culture.

The 'Canton System', which actually evolved during the first sixty years of the eighteenth century, is well known to historians, thanks to the wealth of records left behind by the European trading companies as well as Chinese official documents (Van Dyke 2005). The system was developed in order to gain better control of the activities of Westerners and local officials in the Pearl River Delta, since it was difficult and potentially expensive for the Chinese to exert total control over the comings and goings of boats in the open sea off Macau and its islands, as this would have required extensive patrolling. Forcing boats to sail up the Pearl River into Canton, however, provided the authorities with an ideal passage which was easy to control (Van Dyke 2005). This was organized along typical Chinese lines (Fairbank and Goldman 1998: 195): Chinese merchant families were appointed by the government to act as brokers and superintendents of the foreign traders, with each foreign ship the responsibility of one Chinese family firm. These 'security' merchants formed a guild, the *Cohong*, at the command of a customs superintendent, the *Hoppo*, who answered directly to the Emperor. The *Cohong* and the *Hoppo* assisted, controlled, and taxed the foreign cargoes. This proceeded in three steps: (1) foreign boats were required to call at Macau and were then escorted up the river to Canton; (2) at the harbour of Whampoa (Huangpu), 20 kilometres from Canton, the boats stopped and the cargo was unloaded onto smaller

vessels; (3) from Whampoa, the chief trader of the foreign firm (the *Taipan* or *Supercargo*) and the cargo moved to Canton, while the ship and crew were left at Whampoa. Note that during the eighteenth century, there were no more than 20 *Supercargo*, and very few of them visited the foreign quarter at Canton, since they were allowed to do so only once every three years (Shi 1986; Smith and Matthews 2005).

During navigation, Western captains, sailors, and traders were in contact in particular with three types of people, who attended to their necessities while also acting as controllers on behalf of the authorities: (i) pilots, (ii) compradors, and (iii) linguists. In what follows I examine the contact ecologies that evolved between Westerners and Chinese during the Canton System.

#### 8.2.2.1 *The pilots*

Pilots were often Chinese from Macau or Macanese, because it was easier to find among them people with some knowledge of Portuguese. Therefore, in the beginning of the Canton Trade, Western boats could benefit from having a speaker of Asian-Portuguese onboard. Ships sailing up the river included not only the main players in the region, viz., Portuguese, Dutch, and English, but also Danish, Swedish, Russian, and American boats. Two types of pilots were employed to safeguard the passage: the first group, referred to as ‘outside pilots’ by the Westerners but as ‘fishermen pilots’ by the Chinese, were only responsible for taking the boats into Macau. Once there, the ‘Macau pilot’ took over (Van Dyke 2005: 36). While the former were fishermen without any knowledge of Portuguese, there are abundant documents suggesting that the latter Macau pilots had some knowledge of Portuguese and/or English, which they used to give commands. They were often also literate in written Chinese, a sign of high educational background at the time (Van Dyke 2002, 2005; Bolton 2003). These pilots stayed on the ship throughout the whole journey, guiding the boat all the way up to the harbour of Whampoa, as well as probably taking care of negotiations with smaller boats, merchants, etc. on the way. Once docked, the cargoes were taken care of by the ‘compradors’.

#### 8.2.2.2 *The compradors*

On land there were two types of compradors: ‘factory compradors’, who lived in the trading houses they assisted, and ‘ship compradors’, who supplied the boats but did not spend their entire time on them. The former can thus be said to have been in more intimate contact with Westerners than the latter, though we do know that some compradors went onboard the ships, illegal as this practice was. We also know that some compradors learned the languages of their Western employers, and that it was necessary for all of them to speak a variety of broken English mixed with Portuguese as well as other European elements (Fels 1999; Bolton 2003). Compradors assisted foreigners in providing

provisions for the voyage back, skilled labourers (carpenters, sailors, etc.), and women (and other slave/servants – ‘coolies’), as well as in aiding in cases of accidents, illnesses, etc., and had to be on duty 24 hours a day. In this sense, they were in intimate and intense contact with the foreigners. After the collapse of the Canton System, some of the compradors became independent merchants relying on their knowledge of a Western language, at least CCP, which they had acquired during this phase.

### 8.2.2.3 *The linguists*

Actual business was carried out through the assistance of a third group, the linguists, who were officially nominated mediators between the Western traders and the Chinese authorities, in particular customs officers and governors. Recall that foreigners were forbidden to learn the Chinese language, and any local discovered teaching them would be put to death. Conversely, as Van Dyke (2005: 77–8) notes, it was not advisable for the linguist to be too skilled in a European language, because this could attract the suspicion of the authorities. Therefore, whether skilled or not, it was not in the interest of a linguist to openly display his linguistic skills, and we are in no position to evaluate their real competence in languages such as English or Portuguese. We do know, however, that they had some knowledge of these languages, and it has been suggested that, as mediators in trade between different parties, feigning ignorance or lack of understanding might also have been a recurring strategy to gain advantage and avoid embarrassing situations, especially in encounters between lowly, unrefined Western merchants and the *Hoppo* (Van Dyke 2005).<sup>5</sup> The linguists were in fact responsible, like the compradors, for anything that happened in the European community; they translated during the official trading session and, perhaps more importantly, mediated before the official sessions occurred, in order to ensure that all parties had come to an agreement beforehand.

The first linguists were Macanese, and Makista would have likely been the lingua franca. However, around 1740, a pidginized form of English was also being used between the Chinese and the British (Van Dyke 2005: 77, 81). This shift is likely to have been caused by a combination of factors. Most important is the fact that, as the British gained power in Southern China, Portuguese influence was waning.<sup>6</sup> Towards the end of the eighteenth century, as country trade between India and Macau increased, maritime activities in Macau diminished, to the point that the Macau colony was described as “dying” (Gunn 1996: 29). Also, it was in the interest of Western merchants to depend less on the Macanese interpreters, to avoid being cheated. In addition, there were large Scandinavian<sup>7</sup> companies involved in the trade, as well as Muslim, Armenian, and other traders for whom English represented a more convenient medium of communication. It is noteworthy that throughout the Canton Trade, there were

in fact very few linguists, growing from two to five over the entire period.<sup>8</sup> In time, the Chinese linguists were replaced by foreign missionaries (Van Dyke 2005: 91).<sup>9</sup>

#### 8.2.2.4 *Informal encounters*

In addition to the formal interactions sketched above, informal encounters also took place as part of the Canton Trade. Until the end of the eighteenth century, the crew of the boats had to remain in the port of Whampoa for three months during winter; there could be up to two thousand British sailors at a time as well as other foreign crew, and we know that interactions with the locals (involving, for instance, prostitution and street fights) were common (Bolton 2003; Martino 2003: 51). One of the most notorious sentences in CCP, in fact, in which a rare mixture of Portuguese and English elements is found, certainly comes from this type of environment: *Carei glandi hola pickenini hola?* ‘you want the big whore or the small whore?’ (Noble 1762: 240). It is also interesting that, in the mid-nineteenth century, it was reported that in Whampoa the Chinese could speak better English than in Canton (Nicol 1822; Williams 1836). This implies that, notwithstanding the tight official control, an underworld of material and human smuggling allowed contacts between Westerners and Chinese to take place. This is particularly relevant in the light of the notion of ‘co-habitation societies’ (Faraclas, Corum *et al.* 2007) which stresses the fact that, in reconstructing the linguistic ecology of a contact scenario, formal as well as informal encounters must be taken into account, as well as the relevance of each microecology in the larger picture of evolution of the contact language.

#### 8.2.3 *The nature of early contacts*

The first point to note is that, as may be gleaned from the above, in the Canton System, Chinese–British relations were in fact not symmetrical. Until the First Opium War (1839–42), the British were held in a lower cultural position, due to the general contempt that Chinese had of foreigners, as well as the particular superiority of the Portuguese in linguistic and cultural matters. In addition, the revenues from the trade with the West were marginal to the Chinese compared to the revenues of their Asian trade: Chinese trade with the Philippines, Southeast Asia, Central Asia and Japan was far more profitable (Fairbank 1953). In point of fact, it was exactly because of this overall position of inferiority that the British engaged the Chinese authorities on such a shady issue as the Opium Trade, and used it as an excuse to fight the first of a series of wars leading to ‘unequal treaties’ that forced China to negotiate terms of trade favourable to Western powers. Considering the overall inferior position of the British in this phase, if a

pidgin English had already started developing, it is tempting to follow Hall's (1944) suggestion that it must have been the British who started 'pidginizing' their own language.<sup>10</sup> What is clear is that, in the particular ecology of the Canton Trade, it is plausible that the British – and, later, other Western traders using English as a lingua franca – contributed to the evolution of a pidginized English, at least in the early years, desperate as they were to engage in trade with a superior power who held them in contempt and was very marginally interested in their trade.

Secondly, one must acknowledge the role of local populations in the various microecologies of the Canton Trade. All phases of the Trade offered opportunities for some degree of contact between Asian-Portuguese, English, Cantonese, and perhaps official Chinese (Mandarin). Interactions between pilots, interpreters, and the crew brought Cantonese, Makista, and English into contact. The compradors, despite the limitations imposed on the factories, were agents of contact between English and Cantonese (as well as other Sinitic languages such as Hokkien, Teochew, and Hakka), as observed by Fairbank (1953: 13, my emphasis):

All contact with foreign merchants was mediated through a special class of compradors, linguists, shroffs and Chinese merchants, as well as through a *special language*.

Thirdly, we should not think of the Canton Trade as the only locus of contact: the Opium Trade, which involved a great deal of smuggling, put Western traders and local interpreters, buyers, and dealers into contact in various coastal areas of China (Fairbank 1953). Other aspects of illicit commerce would have had the same effect; Chinese parties also engaged in illegal trade as this would result in much higher profits for themselves. Corruption was indeed so high that eventually it brought the Canton System to a halt (Fairbank and Goldman 1998). The multiplexity of interactions in the early phase of East–West encounters – both within and outside the Canton System – underscores the importance of historically nuanced approaches to language contact, as advocated among others by Faraclas, Corum *et al.* (2007) and their notion of co-habitation societies, in which microecologies of informal contacts are given serious consideration.

I propose thus that CCP emerged as the result of contacts between Cantonese, British, and other Western traders during the Canton Trade as well as in the illegal activities of smuggling and prostitution. Considering that all Western traders, including the British, were held in a lower social position by the Chinese, and bearing in mind the sporadic and often furtive nature of the contacts in the early days of the Canton Trade, Western agency in the formation of CCP must be assumed. The contacts between English and Makista, it seems, were even more superficial and short-lived, which explains why CCP shows very little influence of Makista and Portuguese.

Table 8.1. *Contacts in the Canton region (adapted from Martino 2003: 62)*

| Factories        |              |                    | Whampoa          |            |                    |
|------------------|--------------|--------------------|------------------|------------|--------------------|
| Formal relations | Foreigners   | Informal relations | Formal relations | Foreigners | Informal relations |
| Traders          | Supercargo   | Servants           | Compradors       | Sailors    | Local sailors      |
| Compradors       | Interpreters | Compradors         |                  |            | Prostitutes        |
| Interpreters     | Missionaries | Shopkeepers        |                  |            | Other locals       |
|                  |              | Coolies            |                  |            |                    |

#### 8.2.4 *Expansion and decline*

Between the eighteenth and nineteenth centuries, several political and economic changes took place that resulted in a great influx of trade. The First Opium War (1839–42), fought in order to force Chinese authorities to release monopoly on commerce and to liberalize the Opium Trade under British control, saw the British emerging as victors and forcing China to open a total of five treaty ports (Canton, Amoy, Fuchow, Shanghai, and Ningpo), where the British were free to trade with any Chinese party. It also resulted in Hong Kong becoming part of the British Empire. With further concessions resulting from the Second Opium War (1856–60), Western trade expanded along the China coast as well as inland.

The East India Company was one of the main contexts of English usage in Southern China at the beginning of the nineteenth century, though English was also used by American traders and by the country traders operating in the Chinese hinterland (Bolton 2003). The role and function of the East India Company had already been increasing since the late eighteenth century, due to an expansion of trade and further recognition of its key role by the international community. As a consequence of this expansion, the community of the factories based in the foreign quarter of Canton started growing and social relations increased in complexity. Martino (2003: 62) captures the contacts between Westerners and Chinese in the Canton area as represented in table 8.1 (see also Bolton 2003: 157).

In 1836 the foreign community in Canton counted approximately 200 anglo-phone individuals, the majority of them males, as Western women tended to reside in Macau.<sup>11</sup> An additional 100 individuals, over half of them of Indian or Persian origin, and the rest Western European, predominantly Portuguese, also lived in the foreign quarter (Morse 1926). In 1827 the first English periodical appeared, which is often taken as an indication of the expansion of the foreign community. In the early to mid-nineteenth century, CCP booklets compiled by Chinese for Chinese started appearing and circulating (Li, Matthews and Smith



2005; see section 8.3). The existence of such manuals testifies to the importance of CCP during that period, as well as to the high level of education and the importance of instruction in the Chinese community. In the late eighteenth century, as a consequence of the Opium Wars and certainly also with the help of such booklets, CCP spread to other parts of China (e.g. Shanghai). CCP was reportedly also used among Chinese of different linguistic backgrounds and of course within the non-English-speaking Western community.

As already described, it is quite clear that, despite official controls, there were interactions between Western and local Chinese in and around the foreign quarter of Canton during this period (Hunter 1885). For example, within the factories, merchant–servant interactions brought the Chinese and English into frequent contact. Inside the trading houses, as noted in Selby and Selby (1997), the relationships between merchants and their employees were quite varied: there were not only cooks and cleaners but also accountants often related to the Chinese compradors. These individuals were often of higher social extraction and well educated. It is in fact to one of these key figures that the compilation of a CCP glossary is attributed. Outside the factories, interactions with shopkeepers and coolies took place on an everyday basis. With commerce intensifying, in the port of Whampoa, informal relations between crews and locals must have also sustained a multilingual ecology.

It is in the twentieth century that CCP started declining. This is usually attributed to the impact of English-medium instruction becoming available in missionary schools (Bolton 2003). However, another factor should be taken into account, namely the dissolution of the very ecology that produced CCP in the first place. First, there was the decline of the Canton Trade around 1830, after the East India Company closed their activities there (Van Dyke 2005), although CCP must have continued to be spoken in other trading enclaves, most definitely in Hong Kong. Secondly, World War II and the Japanese occupation caused a dramatic change in the ecology of Chinese coastal areas, in particular Hong Kong (Mühlhäusler and Baker 1996: 518). While in the twentieth century, CCP could be heard in Hong Kong, Shanghai, and Macau amongst shopkeepers, domestic servants, and sailors (Mühlhäusler and Baker 1996: 517), after the Japanese occupation, a new wave of expatriates replaced the community that had been displaced during World War II. Even if knowledge of CCP had survived among shopkeepers and domestic servants, the new expatriates had no knowledge of CCP and this would have led to its decline.

After World War II, Hong Kong rapidly moved towards a modern and cosmopolitan society. The complex, informal ecologies of the Canton System, the illegal trades run by merchants of varying social extraction and with different educational backgrounds, the interaction of different European vernaculars and Chinese languages, and the mediation of more or less competent middlemen all came to an end then. From the 1960s, the population of Hong Kong



converted to a modern, Anglo-Chinese bilingual (and bicultural) identity. The new expatriate contingent spoke various English varieties, all more or less in standardized form. The ecology of CCP had therefore *de facto* dissolved, and with it the *raison d'être* of CCP.

### 8.3 Grammatical features of China Coast Pidgin

Whether CCP should be considered an example of a stable pidgin has been debated (Shi 1991; Bolton 2003: 195); as shown below, the structure of CCP was indeed rather variable. In this section, I pay particular attention to the role of Cantonese in the evolution of CCP. There are basically two views on this: while Shi (1986, 1991) ascribes peculiarities of CCP to either English or Cantonese influence, with a dominance of the latter, Baker and Mühlhäusler (1990) tend to downplay the role of Chinese in favour of ‘universal’ pidginization patterns. Both approaches have strengths and weaknesses. As noted in Bolton (2003: 161–2), Shi relies heavily on the work by Hall (1944) and may contain misinterpretations of the data. On the other hand, Baker and Mühlhäusler (1990) seem to take for granted that universal patterns of pidginization exist, something that the approach presented in chapters 4 and 5 makes controversial. In addition, though they base their analysis on a large corpus, much of the material in that corpus reflects a type of CCP that can be ascribed to Western, rather than Chinese, speakers of CCP (Bolton 2003: 161, 191). The data discussed in this section, on the other hand, come from Chinese sources and reflect, as I argue, a later phase of CCP developed by Chinese for Chinese (see section 8.4.1).

At this point, a brief discussion of the sources of CCP data is in order. Two manuals for the instruction of CCP are known to us today: (1) *The Common Foreign Language of the Redhaired People* (henceforth *Redhaired Glossary*), published in Canton around 1835, and (2) *The Chinese and English Instructor* (henceforth the *Instructor*) from 1862 (Selby and Selby 1997; Bolton 2003; Li, Matthews and Smith 2005). The existence of such booklets had been noted already by Williams in 1836 and later by Leland in 1876. The first of these ‘chapbooks’ consists of just under 400 lexical entries, some of which are short phrases, covering general communication, business terms, etc. It shows similarities to a Portuguese–Chinese phrasebook dating from 1751 found in Macau, discussed in the next section. The second book is a rather more significant work; compiled by a Chinese comprador by the name of Tong King Sing (Li, Matthews and Smith 2005: 79), it was developed primarily as a manual for English instruction, but contains a number of dialogues in what is referred to as the ‘Canton foreign language’, i.e. CCP (Li, Matthews and Smith 2005: 81). Both chapbooks use Chinese characters to capture English pronunciation; this is a common method of Chinese dialectology already adopted to represent

non-standard pronunciation. In the case of CCP, it presents a challenge as, depending on whether the characters are read according to the pronunciation of Cantonese, Mandarin, or other Chinese variety, a different sense emerges of how the English words and phrases were actually read (Smith 2004).

An example of how this works, based on the *Instructor*, is given below (from Li, Matthews and Smith 2005: 82):

|                                |                                    |
|--------------------------------|------------------------------------|
| It is a bad thing to go to law | 哥羅必剪威黎必                            |
| <b>go law pidgin velly bad</b> | <i>gò lòh bìt jín wài làih bìt</i> |

Each Chinese character (top right) represents a syllable of the CCP word to be rendered (bottom right); in this case, a Cantonese reading is clear (as in most cases). The *Instructor* is particularly useful not only because it is rich in grammatical information, but also because it offers a less anglicized version of CCP compared to earlier material collected from Western speakers of CCP (Tryon, Mühlhäusler and Baker 1996: 486; see also Bolton 2003). As this text reveals, Sinitic influence in CCP is much stronger than previous data allowed us to realize, suggesting either a gradually increasing Sinitic influence in CCP's evolution or the existence of different CCPs, one of English and one of Chinese flavour.<sup>12</sup>

### 8.3.1 Words and sounds

CCP words derive mostly from English, with some Portuguese, Malay, Hindi, and, obviously, Cantonese influences. Some examples include (from Martino 2003): *pidgin* < *business* (Eng.); *catchee* < *catch* (Eng.); *two muchy* < *too much* (Eng.); *Joss* < *Deus* (Port. 'God'); *sabee* < *saber* (Port. 'to know'); *chop* < *chapa* (Mal. 'chop'); *tiffin* < (Hind. 'lunch', in Shi 1991: 10); well-known Cantonese words include: *taipan* (Cant. 'supercargo', lit. 'big class'); *fankuei* (Cant. 'Westerner', lit. 'foreign devil'). Note also the use of *houso* 'house' to indicate family, calqued on the Cantonese word *ga* 'house/family'.

Words of clear Portuguese origin are much fewer than one would expect, considering the potential influence that Macau could have exerted in the early days of the Canton Trade: Bolton (2003: 286) identifies 20 to 22 Portuguese words out of 372 in the *Redhaired Glossary*, which suggests an overall limited influence of Makista on CCP, further discussed in section 8.4.

It is difficult to discuss the phonology of CCP for several reasons. As noted above, the transcription system, using Chinese characters, makes it difficult to know what the exact pronunciation of words was. Moreover, different dialectal readings of the characters yield different sounds (see Shi 1991; Li, Matthews and Smith 2005). Considering the methodological difficulties in analysing romanization, as well as the inherent variability to be expected in

phonology, I do not intend to discuss the phonology of CCP here, in particular because it is too early at this stage to have comparative data from the *Instructor* (see Ansaldo, Matthews and Smith to appear). Nonetheless some general observations are the following. One of the most robust features was replacement of /t/ by /l/: since the former does not exist in Cantonese, it is thus replaced by the only non-nasal sonorant of Cantonese (Shi 1991: 14): *rice* > *lice*. Baker and Mühlhäusler (1990) find no significant influence of Cantonese in the vowel system; however, heavy restructuring of syllable structure to avoid consonant (C) clusters is attributable to Cantonese influence, such that CCV > CVCV, e.g. *stop* > *sitap* (Bolton 2003: 162). Also, final consonants are typically deleted since syllable-final non-nasal consonants are not found in Cantonese.

### 8.3.2 Nouns and Noun Phrases

The structure of the NP reveals variation between Sinitic and non-Sinitic constituent order, as well as typical Sinitic features such as the use of the classifiers *pisi* (or *piecee*) and *chop*.

- (1) *Sixty pisi head* (Shi 1991: 20)  
 60 CL head  
 'Sixty heads'
- (2) *Thisee chop tea what name?*<sup>13</sup>  
 DEM CL tea what name  
 'What is the name of this tea?'

Both patterns illustrated above, [NUM-CL-N] and [DEM-CL-N], are typically Sinitic, more specifically Cantonese (as seen in (3)), and are very frequent in CCP.<sup>14</sup>

- (3) *Yāt go yàhn*                      *Nī go yàhn*  
 1 CL man                      DEM CL man  
 'One man'                      'This man'

The pronouns of CCP have attracted some attention in the literature. Besides what appear to be the regular English forms, Baker and Mühlhäusler (1990: 104) and Tryon, Mühlhäusler and Baker (1996: 488) note that three first person singular forms were found: *I*, *my*, and *me*. In a first phase, all three forms can be found in subject position, with the latter two also used in object position. In a later phase only *my* is found in both functions, which could be taken as a sign of stabilization in the system of CCP. Baker and Mühlhäusler (1990) note that this took place rather rapidly only after the 1830s, before which *I* and *me* were normally encountered. They hypothesize that the change was brought about by the introduction of written instructional materials, but, while the timing

strongly suggests this, hard evidence is difficult to pin down. Smith (to appear) shows that, while second and third person forms were almost invariant in the *Instructor*, there was considerable variation in first person forms, with some differentiation of subject and object forms. The emergence of *my* as the canonical first person form is in fact somewhat surprising, considering the almost universal adoption of *mi* in pidgins and Creoles worldwide. It is possible that the character 米, which is pronounced *máih* in Cantonese, was entered in the booklets by someone with the Mandarin pronunciation *mi* in mind. Another confusing factor is the superficial similarity in appearance between the characters 米 and 未, pronounced *mi* in nineteenth-century Cantonese, which could have been confused during the printing process.

Compounding, a ubiquitous feature of Sinitic typology, is frequently noted in CCP; in particular in reference to people and functions, a compound with the form *man* ('man', calqued on Cantonese *yàhn*) is found: *ba baa man* 'barber'; *se leei man* 'sailor'; *guk man* 'cook' (from Martino 2003: 86). Other compounds of the type [Modifier-N] also occur, e.g. *Joss pidgin* 'religion' (lit. 'God business'); *Joss house* 'temple'. Another calque of a Cantonese word is found in the term 'fashion', rendered as *fasi* (Cantonese *yéuhng* 'manner', as in *dím yéuhng* 'what manner' or 'how', Shi 1986), e.g. *wat fa si* 'what fashion' or 'what way'; *niu fa si* 'new fashion' or 'new way' (Martino 2003: 87). In the *Instructor*, *how fashion* (4a) is used in a way that is parallel to *dím yéuhng* (4b) in Cantonese:

(4a) *My savvy how fashion do* (Instructor IV.33)  
'I will know how to act'

(4b) *Ngóh jī dím yéuhng jouh* (Cantonese)  
'I know how fashion do'

### 8.3.3 Copulas, zeros, and existence

At least two different types of copula seem to occur in CCP: (i) 'have', usually realized as *hab* (5 and 6); this can also be used in its lexical function as a possessive verb and as an aspectual marker (5); and (ii) *belong* (7).

(5) *Chinese man very great rogue truly, but have fashion, no can help.*  
'Chinese men are real rogues but that's how it is, can't help it.'  
(Anson 1748 in Baker and Mühlhäusler 1990: 103)

(6) *My hap go court one time* (Instructor IV.4)  
'I have been to court once'

(7) *You belong honest man* (Selby and Selby 1997: 136)  
'You are a honest man'

According to Baker and Mühlhäusler (1990: 103), while the aspectual usage of (i) increased, (ii) took over as a copula. Eventually, *hab* and *habgot* blended in possessive constructions.

Note however that all copula constructions are rare, as, more often than not, CCP uses the zero copula, as shown in (8), and thus follows Sinitic typology, as noted by Hall (1944):

- (8) *Englishman very good talkee; all heart bad, – no talkee true – too much a proudy.*  
(Selby and Selby 1997: 138)

Moreover, as noted in chapter 7 in the discussion of Baba Malay (BM)–Hokkien similarities, existential verbs in Sinitic (e.g. Hokkien *u*) and Malay (*ada*) are typically used to predicate existence:

- (9) *Léih yáuh géi dō go saimān jái?* (Cantonese)  
you have how many CL small child  
‘How many children do you have?’

Existential verbs in Sinitic can also be used to mark aspect and may also function as emphatic markers. Here the parallel with CCP is evident, as *hab* is first and foremost an existential verb. The verb *got* is likewise used in existential phrases, just like the Cantonese verb *yàuh*, which indicates location, possession, and existence (Matthews and Yip 1994), e.g.:

- (10) *No got suchee thing* (Instructor IV.51)  
‘There is no such thing’

Interestingly, the same usage is found in Singlish: *got mango or not?* ‘Are there mangos/do you have any mangos?’

It is possible that the phrase *but have fashion* in (5) above may not be a copula structure at all and should be interpreted as ‘but such things exist’.<sup>15</sup> In the same example, the second phrase is difficult to interpret, because ‘fashion’ can mean many things in CCP and may be translated as ‘way’. Consider the following sentence: *so fashion you buy some beefoo*, ‘well, you better buy some beef’ (Instructor VI.26). Here ‘fashion’ clearly has no lexical nor functional meaning, as is often the case in CCP. In the *Instructor*, all instances of *hap* and *hapgot* follow the Sinitic existential pattern and are thus not clear-cut copulas, as is already evident from Hall’s (1944) data, suggesting strongly that CCP predominantly uses the zero copula.

As far as *belong* is concerned, it is worth noting that, in most occurrences in the *Instructor*, it is used either in its lexical function ‘to belong to’, or to denote possession. Example (7) above can either be read as ‘you are an honest man’ or ‘you belong to the honest ones’. This is in fact similar to the BZM *punya*, the linking particle typical of PDM varieties (see chapters 3 and 7). Recall that *punya* is used as a possessive marker in Bazaar Malay (BZM), as well as for attributive predication, and also functions as a relative marker (see section 7.3.3). In (11) below we see that all three usages are also possible in CCP, although the third is a bit of a stretch. This could either be an independent development or influence from BZM on CCP.

- (11a) *These belong you?* (Instructor IV.53)  
 ‘Is this yours?’
- (11b) *The tea belong first crop* (Instructor VI.14)  
 ‘This is first crop tea’
- (11c) *This belong my plum cashee* (Instructor VI.3)  
 ‘That is what I paid for it’

If BZM was indeed available in the ecology of this region, it might explain why a Portuguese Pidgin was not necessary, as BZM would have served as language of interethnic communication between early Portuguese settlers and local people. In any event, it is unclear whether *belong* really represents a copula or whether *belong* is to be considered a general linking particle attributing a relation between two phrases (see Shi 1991: 24).

Regarding negation, *no* (usually realized as *lo*, as Cantonese does not distinguish between /l/ and /n/) or *no got* is used in CCP, e.g. *missy no got houso* ‘Miss is not home’ (Instructor VI.38); the negative existential is typical of Sinitic, as explained in chapter 7.

#### 8.3.4 Verbs and Verb Phrases

Another reason why I am sceptical about talking of copula in CCP is the fact that adjectival and adverbial phrases behave predicatively, like property verbs, a feature that usually correlates typologically with zero copula, typical of Sinitic grammar, e.g. *léih hóu fàaih* [you very fast] ‘you are very fast’. This is a very frequent correlation as already discussed in chapter 7 and illustrated in (12) to (14):

- (12) *Court expensee too muchee* (Instructor IV.32)  
 ‘The court fees are very heavy’
- (13) *My too muchee trouble* (Instructor IV.32)  
 ‘I was very much bothered’
- (14) *Thisee wine glassee no clean* (Instructor VI.47)  
 ‘This wine glass is not clean’

Interrogative phrases vary: while English-language sources typically show *wh*-phrases fronted as in English, the *Instructor* data show frequent use of *wh*-in situ following Chinese syntax (15 and 16). We also find the typically Sinitic V-not-V pattern, as in *can no can* ‘can you do it?’

- (15) *How muchee more you wantchee?*  
 ‘What more do you want?’
- (16) *You give what price?*  
 ‘What price do you give?’

Another area of variation reflecting differences between the sources compiled by anglophone speakers and the *Instructor* is adverbial modification: Selby

and Selby (1997) note that time and place adverbs follow English syntax, but the *Instructor* offers many examples of pre-verbal adverbs, which are in line with Cantonese syntax, as in *Ngóhdeih tīngyaht būn* [we tomorrow move] ‘we move tomorrow’ (see Matthews and Yip 1994).

- (17) *Tomorrow come catchee money* (Instructor VI.17)  
‘Come get the money tomorrow’
- (18) *He every day tipsy* (Instructor IV.55)  
‘He gets drunk every day’
- (19) *Inside country justee now no got too muchee tea* (Instructor VI.20)  
‘There is not much tea in the country just now’

Perhaps the most obvious Sinitic feature is pre-verbal prepositional phrases (PP) (Matthews and Smith 2005; Ansaldo, Matthews and Smith to appear), as can be seen in the patterns with *(a)long* (20).

- (20) *My (a)long you takee alla* (Instructor VI.8)  
I from you buy all  
‘I will buy the whole from you’

Matthews and Smith (2005) suggest that the preposition *(a)long*, derived from English ‘along (with)’, behaves like Cantonese *tùhng* ‘with’, since it covers the same three functions of Comitative, Benefactive, and Ablative, seen in (21) to (23). These same functions are also found in *(a)long*, seen in (24) to (26).<sup>16</sup>

- (21) *Kéuih tùhng yāt go pàhngyáuh gong* (Matthews and Smith 2005)  
s/he with one CL friend talk  
‘He’s talking with a friend’
- (22) *Ngóh tùhng léih ló yāt go*  
I with you take one CL  
‘I’ll get one for you’
- (23) *Ngóh tùhng léih máaih yéh*  
I with you buy things  
‘I buy things from you’
- (24) *My no long you buy anymore* (Instructor VI.26)  
‘I won’t have anymore to do with you’
- (25) *You can (a)long my catchee one piecee good boy* (Instructor VI.51)  
‘You can get me a good boy’
- (26) *My long he borrow* (Instructor IV.57)  
‘I borrowed from him’

The [PP-V] order, as in (24) to (26), and the variation between pre-verbal and post-verbal PP as in (27) are typical of Sinitic typology and cross-linguistically not too frequent (Matthews and Smith 2005; see also Dryer 2003):

- (27) *Ngóh hái Gáulùhng góbihn jyuh* *Ngóh jyuh hái Gáulùhng góbihn*  
I at Kowloon that-side live I live at Kowloon that-side  
‘I live in Kowloon’

Note the position of *góbihn* 'there' following a locative: this is calqued in CCP with *side* (a very common feature of CCP also retained in Hong Kong English), e.g. *come Sydney side* '(she) comes from Sydney'; *bring that egg come thisee side* 'bring the eggs here'.

Serial verb constructions are frequent in CCP and follow a number of patterns, as described in Matthews and Smith (2005) and Escure (to appear):

- (a) [*Makee* + V], implies a meaning of causation
- (28) *Makee help my* (Instructor IV.42)  
'Help me'
- (29) *What time he makee marry?* (Instructor IV.71)  
'When is he going to marry?'
- (b) [*Look* + *see*] (from Cantonese *tai gin* [look see] 'watch')
- (30) *My wantchee look see counta* (Instructor VI.56)  
'I want to check the accounts'
- (31) *You look see dog no bitee you* (Instructor VI.58)  
'Don't let the dog bite you'

Escure (to appear) also notes the frequency of directional serial verbs based particularly on *go*, *come*, *bring*, *sendee*, and *takee*. Directional serial verbs are also a feature of Sinitic languages, e.g. *daap fóchē làih* [take train come] 'come by train'. SVCs can denote single as well as multiple events in CCP, e.g.:

- (32) *Go makee findee* (Instructor IV.45)  
'Go find it'
- (33) *Bring come here* (Instructor IV.43)  
'Bring it here'
- (34) *What time you sendee tea come?* (Instructor VI.16)  
'When are you going to send the tea?'
- (35) *Catchee one piece man go* (Instructor IV.66)  
'Engage a man to go'
- (36) *You can callum one coolie sendee go my shop. I sendee callum he come*  
'Send a coolie to my shop. I'll send for him' (Anon. ca. 1850: 434)
- (37) *Taipan can sen wun piece coolly come up my sop looksee*  
'Please, sir, to send a servant to my shop for them' (Anon. 1857: 452)

There seems to be only one aspect marker in CCP, the perfective marker *hap* or *hab* (Shi 1991: 23): *My hab sick* 'I am/got sick'; *coolie hap shutee alla window?* 'has the coolie shut all windows?' (Instructor VI.53).

In CCP, *alla* can be used as a resumptive marker following a list of items, just like the equivalent Cantonese *dou*: *two man alla same* 'We are both alike'; *green tea black tea alla hap got* 'I have both green and black tea'; *you houso alla man velly well* 'Is your family all well?' In Cantonese we find: *dàaih ge sūu ge ngóh dōu yáuh* 'big ones, small ones, I all have'.



Reduplication is rather limited in the CCP data analysed so far, but include the well-known *chop-chop* ‘quickly’, *chow-chow* ‘food, eat’, *chín-chín* ‘greetings’ and *lob-lob* ‘have sexual intercourse/go to brothels’ (*lob* < ‘love’). The relative rarity of reduplication is congruent with the fact that in Cantonese reduplication is not as prominent as in Hokkien (and other Min varieties), the contact ecology in which BM, BZM, and Makista emerged.

Finally, we note that CCP discourse structure is clearly Topic–Comment (Matthews and Smith 2005). In the following examples, the topic, in brackets, is the part that is being spoken of in the comment: [*Good cargo*] *how can sellum cheap?* ‘How can good things sell cheap?’ (*Instructor* VI.11); [*Go law*] *pidgin very bad* ‘It is bad for business to go to court’ (*Instructor* IV.32). As already noted in chapter 6, this is yet another aspect typical of Sinitic languages.

### 8.3.5 Summary

It is most likely that further in-depth analysis of the *Instructor* will reveal deeper aspects of CCP grammar, and we will probably see more evidence of Sinitic substrate. The features above should nonetheless suffice to support the claims that the dominant features of CCP are Sinitic, namely (i) [NUM/DEM-CL-NP]; (ii) zero copula; (iii) existential verbs; (iv) property verbs; (v) pre-verbal adverbs and PPs; (vi) serial verb constructions; and (vii) Topic–Comment structure. In the presentation of the data, I have tried to do two things. (1) I show that Sinitic influence is strong in the data in general, perhaps clearer in the *Instructor* than in the English sources. In this sense, I find Shi’s position, and any sensible substratal claim, more informative and therefore generally stronger than a universalist position. The latter is often proven wrong in the face of abundant data, as seen in the case of Hawai’i Creole English (see section 4.2). If other Pacific pidgins have similar or different features, the first task should be to find out where these may come from in the contact ecology, before relegating them to the varied pool of ‘universal’ features. (2) I show that claims of stability, expansion, etc. are risky when applied to limited data sets open to different interpretations. This is crucial for the study of pidgins: if claims based on a variety like CCP – for which we have relatively abundant material – are problematic, the task seems much more difficult when it comes to pidgins for which we have, at times, merely a handful of sentences. This is a general point: solid theories can hardly arise from scanty or absent data.

## 8.4 The missing Makista link?

As noted in section 8.2, in spite of the geographical proximity, it appears that Makista had very little to do with the evolution of CCP. This ‘absence’

of Makista influence has in fact led to a number of theoretical speculations, invoking the presence of a Portuguese pidgin or the relexification of Makista into CCP. However, as observed by Baker and Mühlhäusler (1990), there is no historical evidence for the emergence of a Portuguese pidgin in the region, and any allusion to relexification of Makista features into CCP falls short of explaining the obvious Sinitic element in CCP. Crucially, a pidgin Portuguese would have had no particular reason to develop, since Asian-Portuguese was already a lingua franca and many mixed vernaculars were available in the region (Martino 2003). Nonetheless Bolton (2003: 166–7) notes a similarity between two written sources for CCP, the *Redhaired Glossary* and the *Aomen Jilüe* or ‘Monograph of Macau’, in that they both feature the engraving of a man in Western (Portuguese) attire which look almost identical. Perhaps more importantly, both lists consist of almost the same number of lexical items, 395 and 372, with some of the entries being almost identical. Is this a coincidence or evidence that could help us establish a direct link between Macau and the evolution of CCP? I focus on this in the next section.

#### 8.4.1 *Comparison of Aomen Jilüe and the Redhaired Glossary*

At first glance the two word lists, dating from 1748 and 1835 respectively, appear similar in aspect and purpose. The *Aomen Jilüe*, or ‘Monograph of Macau’ (MM), was written by two Chinese officers, possibly with the help of a third anonymous Cantonese person; it is the most extensive historical record compiled by Chinese authors on Macau (Bawden 1954). It consists of two volumes divided into three sections: the first describes the geography of Macau, the second the defence structures for monitoring the movements of the Macanese, and the third the habits, culture, and trade of the inhabitants of Macau. The last section ends with a word list of 395 items, presented in Chinese characters with Portuguese equivalents in Chinese transcription. These have been recognized as typically Indo-Portuguese lexical items and short phrases (Gomes 1950; Bawden 1954: 114). The majority of the transliterations appear to be based on Cantonese pronunciation of the characters, though in a few cases a Mandarin reading is used (while some are unclear). The word list covers five rubrics: (i) heaven and earth, (ii) men and things, (iii) clothing and food or attire, (iv) implements and numbers or instruments, and (v) common expressions or words of everyday usage.

Considering that this is the standard arrangement of Chinese vocabularies of foreign languages (Bawden 1954), the *Redhaired Glossary* (RG), on the other hand, presents some noticeable differences. Only four sections are found in the RG: (i) numbers used in business, (ii) common expressions referring to people and things, (iii) everyday language, and (iv) food and miscellaneous (Bolton

2003: 170). A specific section on numbers in business is absent from MM, since the MM focusses less on business terms and more on material culture and social relations than the RG. While the RG word list has 93 items in the business section, the MM has only between 20 and 25 business terms distributed across the last two sections. There is also no clear correspondence between MM (v) and RG (ii) and (iii). MM (i) contains over 90 terms related to climate, times of the year, directions, and places while the RG list has only a handful in section (ii). As we would expect, the Indo- and Malayo-Portuguese connection is clear in the MM through words for specific spices, plants, and animals as well as location, while very few such items can be found in the RG. There is also considerable variation in the characters chosen and their transliteration, a logical fact considering that one aims at rendering Portuguese while the other aims at English pronunciation. Bolton (2003) counts 95 shared items out of 372: this is a small percentage considering the similarities of the environments and the needs such word lists represent, and it seems to me that there are more differences than areas of overlap.

These differences, I suggest, reflect differences between the two ecologies: despite the geographical proximity, there was a substantial social and linguistic gap between the Macau and the Canton communities. Indeed, unlike Canton, Macau had a well-established *patoá* in the form of Makista available to local Portuguese descendants as well as Chinese settlers. It would thus be difficult to see why CCP would develop a strong presence in Macau, as foreign traders had been using Asian-Portuguese varieties since the sixteenth century and since Makista was in existence before CCP.

#### 8.4.2 *Makista and CCP*

I submit that the ‘missing’ Makista link has to do with the fact that Makista evolved in a different type of ecology. As already explained in [chapter 7](#), it was a language first and foremost of intra-communal identity, not just a trade lingua franca. The lack of Portuguese in CCP can thus be explained by the fact that, though interaction between the Portuguese and the English occurred in the early days of British trade in South China, the ecologies in which they functioned were basically very different ones, and contacts were sporadic and possibly hostile. Martino (2003) also notes that the speakers who produced CCP were by and large unrefined, uneducated merchants; their language left no traces in Mandarin, unlike the influence it had on Cantonese (and perhaps other dialects). This indicates a clear register differentiation that supports the idea of separate ecologies. In this respect she also points out differences between the situations in which Jesuits came in contact with Mandarins in the seventeenth century: in the particular encounters that involved intellectuals, the exchanges

led to the acquisition of Chinese language and culture by the foreigners, to which the Chinese were open.

## 8.5      Discussion

From the historical discussion presented in [section 8.2](#), and from the data in [section 8.3](#), we can address some of the interpretations that have been offered regarding the formation of CCP.

It is first and foremost questionable whether CCP was the product of encounters between two ‘high’ cultures, as suggested by Baker and Mühlhäusler (1990). Socio-historical observations clearly show that the British and other Western traders were perceived as inferior by the Chinese until the mid-nineteenth century. I have argued that this suggests English agency in the emergence of CCP, which may explain why there is a body of CCP literature that looks more influenced by English than the data of the *Instructor*. The reason why I have underscored the segregation imposed, and partly achieved, by Chinese authorities is not to suggest that contacts between Chinese and Western merchants were not significant; indeed they were, as stated in [section 8.2.2](#). But it is quite clear that, in an early phase, distance was kept between the parties, and that the Chinese were much less interested in the redhaired Western barbarians than the barbarians were in China. Therefore, the idea of CCP as initially pidginized by English speakers (and other foreigners) is a distinct possibility, as already hypothesized by Hall (1944). Van Dyke (2005: 80) reports at least one attestation of CCP being used by Western merchants in Canton as early as 1715, which according to him suggests that it was through merchants’ needs, rather than through the activities of linguists, that Pidgin English developed. In short, it was most likely initiated by sailors and merchants who were not capable of or allowed to acquire Chinese, mostly for informal trade along the lines of Foreigner Talk (see [chapter 9](#)).

However, CCP really grew in usage next to Portuguese only from the mid-eighteenth century onwards, a time that I believe is more realistic for CCP to have really come into being, as also argued by Martino (2003). Not only were Macau (and Makista) still dominant until that point, but commerce in the first half of the eighteenth century was still very restricted. Considering the superficial and short-lived contacts between British and Chinese, further limited by the social distance imposed by the latter, usage of CCP in the early years must have been very limited. However, after the Opium Wars, the situation changed dramatically: commerce increased significantly, particularly in tea and silk; contacts between Westerners and Chinese were no longer restricted to the Canton factories and Whampoa, but took place in several city-ports of the coast; and foreign merchants were then appointed as customs officers, while foreign companies appointed their own Chinese compradors (Fairbank and Goldman 1998: 203). In other words, the networks became more diverse, and increased in number and complexity.

Considering the *Instructor's* data in the evolution of CCP, we must either talk of a later phase of CCP, strongly Sinitic as shown above, or else of two related registers of CCP, a British and a Chinese one. The answer is probably both: CCP probably started as a very limited jargon used by English speakers (and later Western speakers using English as L2) to try to access local services.<sup>17</sup> Afterwards, as the British rose to power, English became an attractive language to the Chinese. This is corroborated by the publication of the chapbooks, and the spread of CCP as a lingua franca among Chinese with the concomitant evolution of Sinitic grammar. Variation is thus likely to have been there all the while, as already suggested by Bolton (2003). It is difficult to determine whether CCP ever stabilized, due to the ultimately small number of speakers and the relatively short lifespan of the ecologies in which it evolved (Tryon, Mühlhäusler and Baker 1996).

Considering that CCP is one of the best documented pidgins we have, there are several valuable lessons to be learned for the study of pidgins and CLF in general. The existence of written documents representative of the varieties of different speakers proves beyond doubt that significant variation must always be assumed in contact ecologies, even more so where multilingualism is high. This means that any generalization in terms of structure and evolution is always constrained and defined by the set of data that is available to us, and should therefore be made with extreme care. Aspects of language such as 'stability' are often a reflection of the data and not inherent to the linguistic system described. Texts, like speakers, present us with individual grammars that vary from one another, even when ascribed to the same variety. As stated in Mufwene (2001), no two ecologies are the same regarding language evolution.

When we compare the cases of Makista and CCP, we observe an interesting difference in their ecologies. In the case of Makista, we saw that it makes perfect sense to think of the variety as a development of Asian-Portuguese, as proven by its history and grammar. Indeed, continuity is an aspect of Asian-Portuguese Creoles in general, as I have noted in terms of the parallels we find between Indian, Southeast Asian, and Makista Portuguese vernaculars. This is understandable considering the common elements that these ecologies show, especially in terms of intermarriage between Portuguese and local populations and, perhaps more importantly, the close contacts across Portuguese settlements in the Asian region. A comparison of this scenario with that of CCP explains why there is no Makista influence on the latter. It also suggests that social and cultural distance, rather than geographical distance, is the actual hindrance to language contact. Though Makista and CCP evolved in geographical proximity, their reciprocal influence is negligible, because of the social distance between their speakers and the differences in their communicative functions. Makista carried cultural and emotive functions of

identity, while CCP's role was restricted to financial exchanges and trading environments. While Makista was the linguistic expression of a novel cultural manifestation, CCP appeared and disappeared in informal transactions between different parties that converged occasionally towards a set of grammatical patterns which are highly variable. This seems to support the thesis that Creoles have generally not evolved directly from pidgins (Alleyne 1971; Chaudenson 1979; Mufwene 2001) and that, even where there may be some relation between a pidgin and a Creole, a pidgin is only one of the many registers in a multilingual ecology (Den Besten 2007) and nobody ever grows up speaking pidgin only (Le Page and Tabouret-Keller 1985). Pidgins as function-specific codes never acquire a homogeneous network with endonormative patterns and external differentiation. Endonormative behaviour, as I have argued in [chapter 7](#), is typical of populations of speakers that have come to identify themselves as one. When these dynamics are lacking, the system remains unstable and limited; this is why, ultimately, pidgins do not survive over a longer period of time.

It is in this sense that I view pidgin ecologies as microclimates: they are well-defined environments with peculiar characteristics that are radically different from their neighbouring ecologies. Therefore, while it is possible that pidgin-like languages may play a role in the ecology of specific contact languages, at the same time, since pidgins belong to different functional domains, the idea that there must be a direct connection to other existing codes in the multilingual environment should not be entertained, and it is even less likely that they may end up replacing these altogether. As we saw in the last two chapters, the ecology of Makista shows no traces of a Portuguese pidgin ancestor; and the existence of CCP never led to the emergence of a China Coast Creole.

## Notes

- 1 Since this chapter describes the historical period known as the 'Canton Trade', I maintain the English name of the city it refers to: Canton, today known as Guangzhou; further, for the sake of clarity throughout the text and, for coherence, I also use the English variants of the geographical areas that are historically and geographically connected to it: Whampoa and the Pearl River.
- 2 Like other Asian civilizations, the Chinese also had a higher sense of personal grooming and sanitation; in other words, they bathed, something that was still quite rare (and reputedly a health hazard) for Western Europeans of the time. Western men are depicted in Chinese sources as "violent and tyrannical and skilled in the use of weapons ... They wear disheveled hair which hangs over their eyebrows" (Chen 1939: 347).
- 3 This section is partly based on Martino (2003), a very significant study of the social and historical evolution of CCP, grounded in the analysis of substantial Western and Chinese historical sources. I was fortunate to be involved in the final stages of this dissertation, written in Italian, as it is to my mind the most complete study of the context of formation of CCP to date.

- 4 As noted by Ostler (2005), until the eighteenth century, Portuguese functioned as a language of diplomacy and trade on the coast of the Indian Ocean, in the Pacific, and all the way to Japan. English only came into the picture later, with the rise to power of the British East India Company.
- 5 While Van Dyke (2005: 89) reports that linguists would have had only basic skills in English (as is often mentioned in Western accounts of that period, see Bolton 2003), he also offers evidence of very developed linguistic skills. A complaint letter dating back to 1796 written by a trader to the *Hoppo*, the customs superintendent, closes as follows: "I am well convinced your Excellency will not suffer this imposition when it is made known to you therefore I humbly pray your Excellency to interfere therein" (Roe 1967: 190). The letter was translated by the linguist into Chinese and delivered to the authorities. Although it is conceivable that the wording may have changed in the process, it is difficult to imagine how a person with only limited English skills would have been able to understand the content of the letter. This, together with the likelihood that linguists might have concealed their abilities, suggests that perhaps it was not lack of skill we are concerned with but rather lack of overt display.
- 6 Though Portuguese trade with other regions, such as the Maluku and the Philippines, was still strong.
- 7 Bolton (2003) finds evidence of lexical influence from Swedish into CCP, e.g. *anka* 'duck'.
- 8 They did, however, avail themselves of assistants and secretaries in their work.
- 9 Although this state of affairs could have occurred because of their limited linguistic skills, it is more likely to have been a consequence of the fact that such linguists were after all Chinese in the service of the authorities (as noted above) and therefore not to be trusted by Western trading companies.
- 10 This does not exclude the possibility that varieties of English introduced by missionaries in the eighteenth century might have existed in China at this stage, though the historical documentation available at this stage makes it difficult to confirm it (cf. Bolton 2003).
- 11 The British presence in Macau was nonetheless very minor and lasted for a short period of time, until the British took over control of Hong Kong in 1842 (Tryon, Mühlhäusler and Baker 1996).
- 12 The role of the *Instructor* in the scholarship on CCP is particularly intriguing. As noted above, this text was compiled by a Chinese for Chinese and therefore offers us the best source of Chinese CCP to this date. Though the text was known and regarded as the most valuable source on CCP (Selby and Selby 1997; see also Bolton 2003: 172), it was not until 2005 that the data in the *Instructor* became available (Li, Matthews and Smith 2005). At this point, it was discovered that the text contained notes written in the margins, in CCP: these are actual phrases, which contain invaluable grammatical information, the analysis of which is still in progress (Ansaldo, Matthews and Smith to appear).
- 13 This represents reconstructed pronunciation following nineteenth-century English sources; see Li, Matthews and Smith (2005). I have already addressed the difficulties involved in representing the CCP system. The data are presented here according to the original transcriptions with, at times, minor modifications for ease of reading.
- 14 Cantonese data are given in the Yale romanization system.

- 15 It may also be the case that *hab* (also *hap/have*) here is really an emphatic marker that has been mistaken for a copula.
- 16 Note also the similarity between CCP and Tok Pisin *long*, which could result from either historical relationship or common grammaticalization paths (Ansaldò, Matthews and Smith to appear).
- 17 Whether they did so intuitively or with the help of some perceived ‘knowledge’ of Chinese is unclear, but since knowledge of Chinese was difficult to obtain, intuition must have been involved in the process.



## 9 Implications, conclusions, and new horizons

---

*Certainty – that is, twisting the objective nature of our perceptions and fitting our ideas to ‘reality’ or ‘truth’ – is a symptom of ignorance or madness...*

Fernando Pessoa

The reader may have realized throughout this book that I maintain some scepticism towards strong theoretical formulations as well as theory-driven data analysis. Arends (2001b: 182) warns us against approaches that overly rely on theoretical considerations, as they are more notable for their capacity to generate discussion “than [for] their contribution to deepening our understanding of [creole] languages as complex empirical phenomena”.<sup>1</sup> In trying to strike a balance between theoretical and empirical discussions, I have attempted to contribute to a growing evolutionary framework for the study of language change. This framework, developed extensively in Croft (2000, 2001) and Mufwene (2001, 2008), requires, in my view, few and simple ideological commitments. In essence, it asks us to view language as a variable, dynamic, and interactive phenomenon, and rests on two basic assumptions of evolutionary biology: a population view of categories (such as languages), and a theory of selection of evolving systems. I hope that the discussion of evolution of contact languages in Asian contexts, in particular in Sri Lanka, Southeast Asia, and Southern China, offers contrasting case studies to the better-known ones that may not only allow us to discover our own implicit theoretical prejudices, but also present new ideas and alternative explanations to established views through evidence of contact language formation (CLF) in non-exceptionalist scenarios.

I must admit that I actually find it challenging to write a conclusion because, after having spent so many pages on a rather wide range of topics, I feel very aware of how much work remains to be done rather than what has been accomplished so far. Nonetheless, in this chapter I recount what I hope this book has contributed to the study of CLF in general and contact languages of Monsoon Asia in particular. Since each chapter has been followed by specific conclusions on CLF advanced there, this chapter discusses the sum of my findings in relation to general linguistic theory, and concludes with some ideas for future research.

## 9.1 Theoretical and methodological implications

At the beginning of this book, I claimed that I was going to answer the following questions:

1. What kinds of ecologies trigger the emergence of new languages?
2. What is the relationship between input languages and the new grammar?
3. How do speakers use languages in multilingual environments?
4. Why do new languages evolve?
5. Which are the theories of language that can be reconciled with the lessons learned from CLF?

The following sections present a general discussion of issues relevant to these questions, based on the case studies and the theories presented so far.

### 9.1.1 *Ecology*

The overall importance of reconstructing the external factors involved in CLF has already been stated in Alleyne (1979: 96):

Since we agree that pidgins and creoles demonstrate so forcefully the influence of social context on language change, language structure, and language usage, we should be concerned, if we wish to reconstruct the early language and language processes of pidgin and creole speakers, with reconstruction of the sociolinguistic situation at the time of its formation, with the structure of the communicative network, and the communicative needs of different sectors of the population involved.

In general, CLF happens under conditions of population movements, usually involving groups who are outsiders to the context in which they operate, often in a position of mediators, whether empowered or not. Four aspects of ecology already pre-empted in [section 1.3](#) have been particularly relevant to socio-historical analysis, namely (i) slavery and manpower, (ii) political economy, (iii) social networks, and (iv) linguistic diversity. These domains of analysis, which help us reconstruct the external and internal ecology of language, are obviously related. Understanding manpower is essential in order to appreciate the heterogeneity of environments such as precolonial Melaka ([section 2.3](#)), for example, where slaves and servants from many different geographical and linguistic origins interact, leading to the formation of contact Malay varieties ([section 3.2](#)); studying populations thus gives us an indication of the degree of linguistic diversity we may expect. Political economy and social networks shed light on the symmetry of relations that hold in a given community and inform us about which linguistic features may prevail in a given ecology; in this respect, for example, it is revealing to see the different linguistic legacies left behind by the Portuguese and the Dutch in Southeast Asia (sections 2.5 and 3.4). The former developed systems that depended heavily on local resources, leaving

behind a number of hybrid ecologies as a result; the latter relied heavily on external resources, and left little linguistic and cultural trace of their presence.

It seems clear to me that these are not factors specific to Asian ecologies, but rather universal factors of human history. Slavery and population movements are also important factors in reconstructing Caribbean ecologies, as shown, for example, in the work of Arends (2001a). Political economy is crucial in comparing language contacts in different settings, as shown in the typology of political economies proposed in Faraclas, Walicek *et al.* (2007). The importance of social networks has long been recognized, as can be seen in the literature on contact languages in which intermarriage and social relations are given serious consideration (Matras and Bakker 2003). I have claimed that we also need to pay attention to the role of endogamy in identity alignment, as shown in chapters 6 and 7, as this underlies the focussing that determines CLF. This is because one of the strongest motivations for the emergence of new languages is actually new identity formation, that is, the formation of a new social group that sets itself aside from the neighbouring ones, as shown in the case of Baba Malay, Makista, and Sri Lanka Malay.

The attention paid to the role of linguistic diversity in the study of CLF follows in the footsteps of early accounts such as those of Alleyne (1979), Baker and Corne (1986), Baker (1993), and Singler (1996), and has been inspired by the studies of Chaudenson (2001) and Mufwene (1996a, 2001). In a quote that indicates a quasi-evolutionary view of language transmission, Alleyne (1979: 95) reminds us that:

It is important to note that bilingualism was already widespread in Africa and that many Africans coming to the New World were already competent in more than one African language. Furthermore, my claim is that the full European languages were heard, and that if any simplification occurred it was not as significant a factor as the transfer and continuity of the structure of the native languages in the acquisition of other languages.

Similarly, multilingual precolonial Asia is so linguistically diverse and so rich in contact varieties that we are forced to recognize the diversity of linguistic input in CLF from the very outset (see chapters 2 and 3). This means acknowledging the role of variation in language evolution.

### 9.1.2 Typology

It is clear that there is a very close relationship between input languages and grammatical output in CLF: the emerging grammar reflects features of the input languages, as argued by Chaudenson (2001). Logically, where two or more languages are involved, *the emerging grammar can never be the sum of all the grammars in contact*, as this would result in massive redundancy. This is why selection is central in CLF, and recombination of features accounts for the differential output in each case, as proposed by Mufwene (2001).

In the case of Sri Lanka Malay, we observe the power of typological congruence (underscored by Corne 1999) that can steer evolution towards one particular profile; this is of course the manifestation of external and internal factors, in particular of frequency patterns (chapter 6), but it is interesting to discuss them in isolation for the purpose of understanding possible contact dynamics. The evolution of Sri Lanka Malay shows, for example, that morphological elaboration happens if it is typologically sustained, offering a case of ‘complexification’, i.e. the evolution of case markers. The importance of considering the full matrix in which evolution takes place can be seen in the role of Sinhala, which cannot be excluded from a proper account of Sri Lanka Malay evolution, as the typological dominance that accounts for the restructuring observed in the NP can only be explained by trilingual alignments and typological gang-up.<sup>2</sup>

It is thus also clear that adding data from typologically diverse environments leads to the reconsideration of our generalizations on patterns of contact-induced change. In addition to morphological elaboration as typologically motivated, we have seen that typology explains zero copula in contact Malay and pidgin varieties as a matter of selection from features available, rather than as an outcome of alleged simplification processes. A similar analysis was presented in Alleyne (1971) for Caribbean Creoles, where zero copula is modelled on the Kwa languages of West Africa. Other critical perspectives on the typology of contact languages have also been offered by the current case studies. In terms of morphological processes, reduplication patterns of the Asian type, manifested in Baba Malay and Makista, offer a systematic functional alternative to inflectional morphology of the Standard Average European type. In terms of syntax, as shown in China Coast Pidgin, *wh*-fronting is not necessarily the rule in contact-induced change (cf. Arends, Muysken and Smith 1994). Similarly SVO need not be the basic word order: as shown in SLM, SOV can be selected in CLF when typology supports it (cf. Muysken and Veenstra 1994). In addition, variation between the alternative constituent orders of the input languages is most often the case in the contact language, assuming that sufficient data to document variation are being recorded.<sup>3</sup>

In addition, what we should also consider when looking at the typologies in contact is the role of frequency, in terms of perceptual and semantic salience in the internal ecology. In the development of case in SLM, the type (and token) frequency of specific grammatical markers determined the likelihood of their being selected from the matrix: while the highly semantically and phonologically salient Dative evolved as a strong feature of the language, the overall less frequent Accusative became only a dispensable and variable feature of the grammar.

### 9.1.3 *Input–output relations*

Givón (1979b) argues against the notion of simplification in Creole formation on typological grounds: he points out that the idea of simplification as

observed in Creole formation is likely to derive from the fact that West African languages are at the analytic end of Sapir's typology scale, a point reiterated in Ansaldo and Nordhoff (2009) in the light of the lack of simplification found in a different typological setting such as Sri Lanka Malay. It is clear that isolating typology is heavily involved in the typological matrix of contact Malay varieties, while morphologically rich registers are either rare or absent, which is why I argue against 'pidginization' as the evolutionary route to Pidgin-derived Malay varieties (section 3.2).

In addition to considering the *type* of languages involved in CLF, it is also important to clarify the *variety* of language actually available to speakers in the early phases of CLF, already discussed by Chaudenson (2001). While in some contexts the 'full' language may have been used, in others we find predominantly contact varieties in contact, as has been shown in chapter 7; there I argued that a careful consideration of the variety of Malay that provided the input for a number of contact languages of the region weakens the argument of morphological reduction in evolution, because 'full' varieties were not involved. This means that the case studies presented here offer little support for universal strategies of simplification.

Following from these considerations, I have suggested that, where some reduction may be found, i.e. in the early phases of Asian-Portuguese evolution (chapter 7) and in China Coast Pidgin (CCP, chapter 8), there is substantial theoretical and historical evidence in favour of simplification in the form of Foreigner Talk (FT) strategies, i.e. not an outcome of imperfect learning or distant target, but rather a conscious process of reduction, omitting what may be perceived as redundant features from the communicative point of view (see also Thomason and Kaufman 1988). The conditions that lead to FT are by no means mysterious: Spolsky (1989) synthesizes these as follows: (a) simplifying condition for FT: speaker simplifies his register when he perceives that the interlocutor does not share the same variety; (b) matching condition for FT: great shared knowledge leads to short, simple expressions (and the converse); (c) solidarity condition for FT: preference to assume shared knowledge for solidarity purpose leads to short/simple talk; (d) distortion condition for FT: simplification can be pursued to the point of distortion. In microclimates such as the Canton Trade (chapter 8), where CCP evolved, it seems quite likely that FT might account for the early phases of pidgin development, as the British and other Western traders were marginalized by the Chinese, who were interested in keeping them at a safe distance. While it is clear that Chinese agency was strong in the later development of CCP, as shown in the robust Cantonese substrate features found in the grammar, British agency in the early phases is very likely, leading to a simplification of their own language in order to gain access to Chinese traders. In other words, the input was already, to a high degree, simplified.

In similar vein, we need to recall the suggestion that the variety of Portuguese present in the (Asian) colonies was not a literary or high variety

of Portuguese. In Macanese literature, we find frequent references to the ‘old’ or ‘colonial’ dialect used by Portuguese traders, usually depicted as crude, simple, or unrefined (Ferreira 1983; Batalha 1994). We also find very specific claims regarding the development of a simplified Portuguese created by the Portuguese in Europe and then brought over to their colonies in Africa and Asia (Holm 1989: 269–72; [chapters 3 and 7](#) of this volume). I have already discussed how the basic features of CCP may be attributed to a predominant selection of Sinitic features and how ‘simplification’ was mostly led by Western users under FT conditions. In the case of Makista, the apparent ‘lack’ of Portuguese morphology can likewise not merely be attributed to a universal process of simplification. As shown in [chapter 7](#), plurality, for example, is expressed by reduplication rather than the plural morpheme; tense morphology is replaced by aspectual categories; and zero copula is predictable given the predicative nature of ‘adjectives’ or property items. These are all typological features derived from either Malay or Sinitic and are explained by the nature of the typological matrix in which Portuguese contact varieties came into contact with Asian grammars.

#### 9.1.4 *Simplification revisited*

It is interesting thus to recognize that in the history of Asian-Portuguese and CCP alike, the often-invoked process of simplification can be accounted for along two lines: (a) convergence to isolating typology, partly based on frequency, and (b) reduction of communicatively unnecessary features by speakers of the language (along the lines of FT theory). It is interesting because, if this is correct, we can finally dismiss the correlation between contact, pidginization (and even creolization, if we have not done so already), and impediments, shortcomings or otherwise missing elements in second language acquisition (SLA). In this sense, pidginization, if present, is not SLA under social and psychological distance, which results in partial learning of the target (Schumann 1978b); rather, it is the outcome of speakers adjusting their speech to accommodate their hearers in response to the mutuality they perceive between themselves and their hearers (Ferguson 1975). It is, after all, part of the linguistic competence of adult speakers of a language to know how to simplify it, if social circumstances should warrant this, as reflected in the general thrust of much of Charles Ferguson’s (1982) work. The restriction of words to generic terms, the elimination of allomorphy, the use of parataxis rather than hypotaxis, and the adoption of invariant word order for grammatical relations all ease the processing load of the hearer; there is no need to cope with variations and irregularities. As a communicative strategy used by competent speakers in communicating with those who are not expected to be fluent in that language, FT also resembles another type of simplified speech known as ‘baby talk’

(Ferguson 1971). Such simplificatory strategies are in fact not a universal phenomenon of communication but must be considered culture-specific linguistic practices, as Ochs and Schieffelin (1984) have shown, and as also discussed in section 7.1. It is intriguing to reflect on the possibility that, if FT/baby talk routines account for a substantial amount of the restructuring that the colonial lexifier languages underwent in a number of contact environments, and if these are typical practices of Western cultures, then the agents of the often-noted simplification were most likely the native speakers of the lexifier languages themselves, i.e. Westerners, as already suggested in Ferguson (1971). As already proposed for the evolution of CCP, FT routines are, of course, not invoked as the exclusive route to pidgin formation, since the role of substrate typology is in all instances a fundamental aspect of language evolution. But it seems appropriate to consider FT routines as part of the specific formation of CCP based on socio-historical information, as well as the cause of some of the so-called universal simplification strategies.<sup>4</sup>

The SLA = simplification view, which is reiterated in McWhorter (2007) for the evolution of Malay, for example, simply does not work because such a view is caught in a linear view of transmission. It is not the case that people simplify in SLA scenarios, which McWhorter suggests can explain the evolution of a language such as Malay, which was a lingua franca for millennia. It is rather the case that Malay evolved from a matrix in which isolating grammars were prevalent. It is further the case that speakers of Western languages, at least, simplified their own language while interacting with others. This is related to the fact that, in informal, untutored, multilingual environments, selection and replication of features that are frequent in token and type is more likely, as a result of phonetic prominence as well as semantic and pragmatic salience within the matrix. Note in this respect that inflection, often the showcase of simplificatory approaches (Kusters 1996; McWhorter 2007), has few of these traits, being often employed in functions that are semantically redundant and phonetically reduced (though this varies from one language to another, see Aboh and Ansaldo 2007). This suggests that what speakers are not heavily exposed to, speakers do not easily replicate.

In sum, the evolutionary patterns we find in the case studies presented here demonstrate the role of congruence and frequency in CLF, but do not offer strong support in favour of simplification as SLA in language evolution. The notion of pidginization appears largely problematic as an explanation because any viable notion of simplicity will be either theory internal or language specific (Ansaldo and Matthews 2007). In its strong version, the notion of simplification is most likely a product of a monolithic, linear view of acquisition, as explained in chapters 4 and 5. This ‘monolithic’ view of language is false according to Haugen (1972: 335); it is a useful fiction that however fails to capture the complexities of the ecology of language. In an ecologically driven scenario,



speakers construct their language throughout life, based on the diverse input they receive from their environment (Mufwene 2001; Tomasello 2003), which may indeed change dynamically over time (Lim 2007, 2008, forthcoming).

### 9.1.5 *Alignment*

I have argued in [chapter 7](#) that CLF is the result of identity alignment, i.e. the construction of a new identity in a newly emerging community. This means a view of speakers as agents who actively engage in accommodation and differentiation through language use. This view is also akin to the notion of acts of identity that speakers engage in as a consequence of their desire to identify more with one social group (Le Page and Tabouret-Keller 1985), and the notion of cultural hybridity (Chaudenson 2001). Going beyond repertoires and choices at the level of the individual, the situation of the Malays of Sri Lanka, for example, can be seen in terms of the exploitation within a multilingual community of the codes they have at hand, in order to align their identity according to changing environments and times. It has been suggested that identity options offered to individuals at a given moment in history are subject to shifting language ideologies that legitimize and value particular identities more than others (Pavlenko and Blackledge 2004).<sup>5</sup> I have argued that the evolution of Sri Lanka Malay was the outcome of a similar process of identity alignment, one in which a heterogeneous and diffuse Malay/Indonesian identity was traded for a specific Malay/Indonesian Sri Lankan identity. The convergence to Lankan grammars – in the sense of adaptation to the dominant linguistic context (see Enfield 2008) – together with the retention of substantial lexical items, resulted in the communities achieving internal cohesion, clear self-representation, and identification and integration within their country of adoption.

Where identity is concerned, moreover, Lim and Ansaldo (2007) argue that the conceptualization of the identity of a displaced community as less coherent and robust does not in fact apply to all multilingual communities. As suggested in Lim and Ansaldo (2007), “with communities such as the Sri Lanka Malays for whom multilingualism in a multicultural context is nature – cf. the Western monolingual, one-nation–one-language ideologies that have been imposed on scholarship – we can certainly go beyond the idea of language shift and identity loss, and [perhaps] even beyond the dynamic of continuously constructing one’s own identity through language”. This requires a paradigm shift in our conceptualization of multilingual communities, which should be seen as being multilingual in terms of linguistic repertoire and pluralist in terms of identity.<sup>6</sup> Their negotiations are thus not necessarily a matter of language shift but rather a matter of ‘choice’ of a linguistic resource from their repertoire appropriate for a specific circumstance, and the concurrent alignment with a particular facet of their identity (Lim and Ansaldo 2007).



Such a view is congruent with the notion of linguistic citizenship (LC) as proposed in Stroud (2000) and Heugh and Stroud (2004), where language is seen not as a discrete construct that is generated and owned by speakers and that can be lost by a community, but rather as a semiotic resource for the construction of agency and self-representation, an economic resource and site of political and economic struggle, a global resource to address local–global concerns, and an intimate resource as the foundation of respect for difference on a global level. In the LC paradigm, the community is *served* by its linguistic resources – which comprise pluralist alternatives reflecting the reality of the linguistic market, and, consequently, negotiable multiple, diverse and shifting identities – and is not restrained by its language. Individuals and speech communities choose to empower themselves in what they see as the best possible way with regard to existing power relationships. Thinking in terms of identity alignment also means that for communities such as the Sri Lanka Malays for whom multilingualism in a multicultural context is their nature (cf. the Western monolingual ideologies discussed in sections 4.1 and 4.2), we need to move beyond the notions of shift and death. Therefore, in multilingual communities, individuals do not clearly identify with one specific variety (Lim and Ansaldo 2007); they must identify with the various linguistic codes they have at their disposal *as well as the possibility of mixing them in different contexts and for different practices*.

### 9.1.6 Evolution and language contact

The evolution of contact Malay varieties discussed in particular in chapter 7 suggests that distinguishing between contact-induced change and other routes to speciation (e.g. ‘internal’ changes) is in practice a very difficult issue (which also leads to controversial positions in classifying different languages). In an evolutionary framework, however, as in the approach currently adopted, the issue of speciation and classification receives a novel treatment, as shown in Mufwene (2001: 98), who breaks down the traditional dichotomy between pure and hybrid with the notion of “blending inheritance”; this notion aptly captures the evolutionary path of the contact Malay varieties described in this book. In this view, all languages are mixed, as any acquisition involves the recombination of features from diverse idiolects or inputs – regardless of whether or not they belong to the same variety, as discussed in chapter 4. The idea of multiple ancestry and contact-induced influence applies to the evolution of all languages and questions notions of ‘direct’ transmission and ‘native’ acquisition (Mufwene 2001, 2008). Not only does this view allow us to move away from taxonomies that are necessarily biased historically and ideologically, it also allows us to more closely reflect the historical descent of the languages we discuss, which ultimately increases our understanding of speciation. For

example, we have seen that Baba Malay, Bazaar Malay, and Cocos Malay are all Malay varieties of multiple parents, clearly Malay–Sinitic in the first two cases (and possibly Malay/Indonesian–PDM in the latter case).

The evolutionary framework, moreover, questions the linear, ‘pure’ type of speciation implied in the *Stammbaum* model, related to the critique of linear transmission. In the history of any language, and thus in the history of Malay, we have seen that there is good reason to believe that abundant horizontal transmission took place. Recall also that Adelaar and Prentice (1996) offer a classification of Malay that is socio-historical, as clearly reiterated in Adelaar (2005: 203). This means that the classification is based on domains of usage, not successive historical phases. Consider further that the ecology of Malay is characterized by multilingualism and heteroglossia, from ancient times to the present, and that considerable social distance was found between colloquial and literary varieties. In an evolutionary take on speciation, Malay varieties can thus be seen largely as the outcome of contact-induced change involving different degrees of typological diversity in the typological matrix. As shown in chapter 7, a number of Malay varieties evolved in precolonial and colonial settings out of the contact of Old Malay varieties with other languages. This explains the different typological profile as well as the shared lexical and phonological features of many of these varieties. Following from the ideological stance taken in this book (chapters 4 and 5), this, crucially, does *not* suggest that Malay varieties evolved in a qualitatively different way from, say, European languages; it means, instead, that the possibility that European varieties evolved predominantly through language contact must be seriously considered (see also Mufwene 2001).

## 9.2 Conclusions and new horizons

The theoretical stances of the approach presented in this book, as discussed in chapters 4 and 5, combine essentially three strands: (1) a non-exceptionalist approach to CLF (DeGraff 2003), i.e. one that makes no assumption of extraordinary cognitive, acquisitional, or structural events in CLF, based on an evolutionary scenario in which ecology plays a major role (Mufwene 2001); (2) a functional–typological approach to language (Croft 2000); and (3) a commitment to socially plausible theories of language use. I believe that the study of CLF in the Asian contexts presented in this book has shown that there is nothing exceptional in the evolution of languages such as Sri Lanka Malay, Baba Malay, or Makista. There were no traumas involved that might lead one to suspect cognitive anomalies, there was no lack of fully fledged linguistic systems that speakers were exposed to, and there were socio-historical conditions common in precolonial and colonial times. We have seen that a typologically based approach results in a more sophisticated take on the notions of

simplicity and complexity central to theories of language evolution. Moreover, the functionalist, usage-based analysis is by definition more sensitive to ecology because it takes context of use seriously, and it integrates well with an evolutionary theory of adaptive systems, as well as with the role that frequency, among other factors, can play in a population view of language as constructed by its speakers.

It is notoriously difficult to reconstruct the linguistic history of past contact ecologies: population movements are difficult to record and, even when historical documentation is available, it is seldom possible to ascertain the linguistic origins of displaced groups. If linguistic documentation is available, in the form of word lists, phrase books, and metalinguistic comments, it has to be treated with utmost care, as it is often the product of (non-)linguists who might bring cultural and linguistic biases into the description. However, if we accept the fundamental socio-historical nature of the study of CLF, we must also accept that in order to produce plausible accounts of it, we must engage in historical analysis. Considering the growing amount of historical documentation based on primary sources, there is hardly any excuse not to do so, and the position that ‘history should be left to the historian’ is only acceptable if we also agree that ‘contact linguistics should be left to the historically minded linguist’. In an ecological approach, there is by definition no opting out of the detailed analysis of external factors. This should be a critical analysis which does not rely solely on interpretations and numbers available in the literature. In the case of Sri Lanka Malay, for example, oral history was crucial in re-assessing some of the ideologically biased claims of its historiography – history can after all also be biased. A clear discussion of reconstructive methodology, as offered by Rudolph (1998: 13–22), is summarized as follows: (i) participant observation (when possible); (ii) narratives and expert interview; (iii) text and discourse analysis; and (iv) etymology and conceptual history. This is a highly qualitative method, within which quantitative data analysis can be integrated, as long as one does not succumb to the number game, as numbers can also lie. Besides primary sources (oral history and period accounts), journal articles, popular literature, interviews, and participant observation (if possible) should be used in reconstructing the ecology of a contact environment.

For the study and theorizing of CLF, a number of relevant reflections arise from this study with significant methodological implications:

1. Contact ecologies are by definition multilingual: we must always expect at least two, but more often several, languages to have been in contact where CLF takes place. Mere bilingualism – i.e. the involvement of just two languages – is hard to prove in most contact environments (e.g. Bakker and Mous 1994; Matras and Bakker 2003); as the case of the Babas shows, a high degree of societal and individual multilingualism is often found.

What this suggests regarding the notion of transmission is that, in most contact settings, features transfer from the input varieties to the new grammar through various stages of adaptation and not directly from the source to the product (see Siegel 1997).

2. Language usage largely involves selection and replication of features available from the ecology. The replication also involves spontaneous innovative (or creative) activity in linguistically diverse environments, offering a higher potential for the evolution of new forms, as CLF reveals. Identical and innovative replication corresponds to different manifestations of accommodation as the basic linguistic behaviour of humans, i.e. assimilation, maintenance, and adaptation of features (see Giles 1980; Beebe and Giles 1984). This highlights the importance of a population view of language as a pool of variable features.
3. Typological input is the dominant force in selection and replication. This suggests that analysis of adstrate grammars is fundamental and that, in the light of the typological matrix, it should be as comprehensive as possible. Moreover, as has often been remarked upon, the adstrates used for analysis should correspond to the language varieties actually spoken by the populations that came in contact, typically colloquial, non-standard vernaculars, as well as contact varieties (Alleyne 1979; Mufwene 2001).<sup>7</sup>
4. Following the notion of identity alignment, speakers are creative users of language: the agency in CLF rests with both adults and children, as argued by DeGraff (1999). Therefore we must consider the functions of different codes in multilingual ecologies in terms of adults' and children's usage and focus on synchronic and diachronic variation in our approaches.
5. Transmission from the point of view of alignment involves socialization and links linguistic analysis of CLF to cultural analysis. Non-linguistic aspects of contact ecologies are important pieces of the puzzle in reconstructing CLF (Chaudenson 2001).
6. Transmission is always multiple in contact ecologies, and unilinear evolution is unlikely (see section 4.2). As shown in the cases of Sri Lanka Malay, Baba Malay, and Makista, we must recognize that contact ecologies are characterized by continuous evolution of contact varieties which reflects multilingual transmission and usage.
7. It is unlikely that universals of SLA along the lines of pidginization exist. It is clear that simplification does not always happen in multilingual SLA encounters and that the role of isolating morphology in the matrix accounts for what looks like simplification. To prove that simplification exists, one must look not at multilingual ecologies and cases of restructuring but rather at monolingual ones and non-innovative transmission. It is here that, if we were able to find that between generations a simpler system had emerged, we might locate 'real' simplification. Moreover, if simplification is captured

as FT conditions, it is a native strategy and not a deficiency on the part of learners.

8. Pidgins may exist in multilingual ecologies but are not necessarily directly linked to CLF. As we saw in the case of China Coast Pidgin, what happens in pidgin ecologies is not necessarily related to the evolution of contact languages in their proximity. Surprisingly, we should have known this all along. Alleyne (1979: 104) is correct in pointing out that:

the development cycle of pidgin → Creole → post-Creole characterized by reduction in the pidgin stage followed by continuous expansion into the Creole and post-Creole stages has not been established, and that, certainly for the Caribbean, a general framework of language contact producing different kinds of language change arising out of different experiences of language contact [i.e. different ecologies] can account for all types of language phenomena, including language situations and language planning programs.

9. Abrupt nativization of a restricted code is not attested in the case studies presented here and, as far as I am aware, nowhere else in the region (see also Mühlhäusler 1986; Holm 1989: 631). Since contact ecologies are by definition multilingual, other registers besides a pidgin are always available to children (see Roberts 1998, 2000 for Hawaiian Creole English and Den Besten 2007 for Afrikaans). From the points spelt out in 4 to 6, CLF is a gradual process, as it requires several generations of interaction in order to consolidate.
10. Contact between different idiolects, varieties, and languages is everywhere (Croft 2000; Mufwene 2001); internal change is a mirage arising because of lack of ecological reconstruction. Relying on internal change should be used as a last resort and be considered valid only until sufficient information is revealed to fill the gap in our knowledge.

Following on this, my wish is that more research on contact languages will embrace a holistic approach to the study of CLF, an approach in which the historical, social, and linguistic dimensions come together to produce realistic analyses. There should be a balance of structural and contextual analysis because either of them, in isolation, cannot make any solid claims of general linguistic relevance. More specifically, we need more studies of typologically diverse encounters, beyond the domain of European languages, as well as examples of more varied socio-historical settings, from different geographical areas and covering different historical periods. We also need a re-evaluation of better-known cases in the light of (a) new historical documentation and (b) typological advances in our understanding of substrate features, which allow us to look at input–output relations with a more informed and sophisticated mind than previously possible. In particular, three aspects of the study of multilingualism, it seems to me, should be primary on the agenda of contact

linguists: (1) language contact in trilingual ecologies; (2) longitudinal studies of language socialization; and (3) language choice at the communal and individual level.

Finally, in a linguistic enterprise that aims, among other things, to explain speciation, i.e. the evolution of linguistic diversity, we can posit a range of different ecologies in which evolution could have taken place. I propose that we envisage this as a continuum between two poles, illustrated in [figure 9.1](#).

We can easily imagine that human communities throughout history would have fallen between these two points, but the crucial question is: which pole would have been more common, frequent, or unmarked? Clearly it is ecology A, because ecology B, in its extreme form, is the product of Western European nation states of the last two hundred years and does not account for earlier human settlements in general, whether hunter gatherers or farming communities. The evidence supporting this is clear. As already discussed in [chapters 4 and 5](#), variation, admixture, informal transmission, and creativity are common traits of human communication in many contemporary societies. Even more importantly, Cavalli-Sforza (2000: 202) argues that “there is almost always some genetic exchange between geographic, ethnic or socio-economic groups” in the history of populations, an exchange that implies the ubiquity of language contact. Johnson and Earle (2000) also demonstrate that intercommunity networks and patterns of exchange (as well as warfare) between different human groups have been involved in all social types above the family group in the evolution of human society, be they local groups or regional polities. Moreover, even at the family/camp level, interaction with other groups is likely, as such groups are seldom completely isolated (Johnson and Earle 2000: 32–7). This means that language contact and the possibility of borrowing and interference is ubiquitous in human societies (see also Croft 2003; Mufwene 2001, 2008).<sup>8</sup> This is an important realization because it tells us that, in the evolution

| <u>Ecology A</u>      | <u>Ecology B</u> |
|-----------------------|------------------|
| Multilingualism       | Monolingualism   |
| Language admixture    | Language purism  |
| Informal transmission | Schooling        |
| High variation        | Normativity      |
| Creativity            | Conformity       |

Figure 9.1 Multilingual and monolingual ecologies

of languages, contact, multilingual transmission, and innovation are normal activities that speakers engage in. Therefore, though clearly beyond the scope of this book, I dare offer the following suggestion: that we start considering *monolingual acquisition under normative pressure* as actually rather exotic in the history of human language evolution, and the lessons derived from the study of acquisition and language change in monolingual ecologies as potentially exceptional, maybe even peripheral to the construction of general theories of language.<sup>9</sup> As has been the goal of this book and others before it, this should lead us to seriously question many of our current assumptions of what is normal, universal, or unmarked in the study of language change.

## Notes

- 1 In less guarded words, Feyerabend (1975) claimed that science is in fact much closer to myth than scientific philosophy is prepared to admit, based on the circularity inherent in the theory–data relationship. He reminds us that theories typically generate the types of data that are relevant to them, and it should come as no surprise then that such data indeed prove the theory right.
- 2 Another case study of ‘complexification’ can be found in Dreyfuss and Oka (1979): the authors show that the Chinese Indonesian mixed language spoken in Java, in communities of Javanese women and Chinese men, takes 80 per cent of its grammar from Javanese and 80 per cent of its lexicon from Indonesian Malay; it thus combines all morphological material of Javanese with the BZM spoken by the males.
- 3 These are just some highlights of the data presented in this book, and a full evaluation of their significance for non-Asian contexts must be left to other, future works.
- 4 This is not meant to suggest that Europeans used FT for all their communicative purposes, but that varieties influenced by FT strategies were most likely in contact with non-European varieties (cf. Alleyne 1971).
- 5 In this respect, this postcolonial era affords a globalized existence where formerly displaced communities do indeed have the option and the capacity to reconnect with their precolonial origins, as the Sri Lanka Malays have done with Malaysia, and we can see how the ideology of being included in the larger Malay world is legitimized and valued.
- 6 Djité (2006) notes how multilingual speakers have an increased capacity to communicate across regional and national boundaries through languages that were once limited to one territory.
- 7 Though in the case of indigenized varieties of English, the targets were also scholastic, not vernacular varieties (Mufwene 2001).
- 8 In fact, there is good reason to believe that language change almost always implies language contact; a correlation between isolated populations and lack of linguistic mutation can be observed in cases such as Iceland and Sardinia, where prolonged isolation from contact with other populations led to the maintenance of archaic varieties (Cavalli-Sforza 2000: 202–3).
- 9 The question that arises here is how we should view monolingual behaviour, in terms of usage, transmission, and theoretical relevance. What kinds of ecologies are

constructed in order to force speakers to abstain from the natural creativity inherent in language use? What kind of cognitive pressures do monolingual individuals sustain during childhood, through enforced normativity at school and in society at large, aimed at constraining their natural linguistic behaviour? Why have theories derived from monolingual inquiry dominated linguistics so far? Is there a connection between this dominance and the fact that monolingualism is taking over as an educational paradigm worldwide, contributing to the demise of linguistic and thus cultural diversity? I am convinced that, in discussing these questions, we can make useful contributions to the appreciation of the role of language in human society.



## References

---

- Aarsleff, Hans 1982. *From Locke to Saussure: Essays on the Study of Language and Intellectual History*. Minneapolis: University of Minnesota Press.
- Aboh, Enoch O. and Umberto Ansaldo 2007. The role of typology in language creation: A descriptive take. In Ansaldo, Matthews and Lim (eds.), 39–66.
- Aboh, Enoch O. and Norval Smith (eds.) 2009. *Complex Processes in New Languages*. Creole Language Library. Amsterdam/Philadelphia: John Benjamins.
- Adelaar, Karl Alexander 1991. Some notes on the origin of Sri Lankan Malay. In Heinrich Steinhauer (ed.), *Papers in Austronesian Linguistics*, no. 1. Pacific Linguistics A-81. Canberra: The Australian National University, 1–22.
- Adelaar, Karl Alexander 1994. Malay and Javanese loanwords in Malagasy, Tagalog and Siraya (Formosa). *Bijdragen tot de Taal-, Land- en Volkenkunde* **150**.1: 50–65.
- Adelaar, Karl Alexander 1996a. Contact languages in Indonesian and Malaysia other than Malay. In Stephen A. Wurm, Peter Mühlhäusler and Darrel T. Tryon (eds.), *Atlas of Languages of Intercultural Communication in the Pacific, Asia and the Americas*. Berlin/New York: Mouton de Gruyter, 695–712.
- Adelaar, Karl Alexander 1996b. Malay in the Cocos (Keeling) Islands. In Bernd Nothofer (ed.), *Reconstruction, Classification, Description. Festschrift in Honor of Isidore Dyen*. Hamburg: Abera Verlag, 167–98.
- Adelaar, Karl Alexander 2005. Structural diversity in the Malayic subgroup. In Karl Alexander Adelaar and Nikolaus P. Himmelmann (eds.), *The Austronesian Languages of South East Asia and Madagascar*. Routledge Language Family Series. London: Routledge Curzon, 202–26.
- Adelaar, Karl Alexander and David John Prentice 1996. Malay: Its history, role and spread. In Stephen A. Wurm, Peter Mühlhäusler and Darrel T. Tryon (eds.), *Atlas of Languages of Intercultural Communication in the Pacific, Asia and the Americas*. Berlin/New York: Mouton de Gruyter, 673–93.
- Aikhenvald, Alexandra and R. M. W. Dixon (eds.) 2001. *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*. Oxford University Press.
- Aikhenvald, Alexandra and R. M. W. Dixon (eds.) 2006. *Grammars in Contact: A Cross-Linguistic Typology*. Oxford University Press.
- Albuquerque, de Braz 1557. *The Commentaries of the Great Afonso Dalboquerque*. Transl. by W. de Gray Birch 1877–1880. London: Hakluyt Society.
- Alleyne, Mervin 1971. Acculturation and the cultural matrix of creolization. In Hymes (ed.), 169–86.
- Alleyne, Mervin 1979. On the genesis of languages. In Kenneth C. Hill (ed.), *The Genesis of Language*. Ann Arbor: Karoma, 89–107.

- Amaro, Anna M. 1994. Sons and daughters of the soil: The first decade of Luso Chinese diplomacy. *Review of Culture* 20: 12–67.
- Andersen, Henning 2006. Synchrony, diachrony and evolution. In Ole Nedergaard Thomsen (ed.), *Competing Models of Linguistic Change: Evolution and Beyond*. Amsterdam/Philadelphia: John Benjamins, 59–90.
- Annamalai, E. 2001. *Managing Multilingualism in India: Political and Linguistic Manifestations*. New Delhi: Sage.
- Anon. 1833. Description of the city of Canton. *Chinese Repository* 2.7: 296–306.
- Anon. ca. 1850. *Hühng mōuh tung yuhng fāan wah* [The Language of the Redhaired Foreigners]. Guangzhou.
- Anon. 1857. Canton-English. *Households Words* 15: 450–2.
- Ansaldò, Umberto 1999. *Comparative Constructions in Sinitic: Areal Typology and Grammaticalization Patterns*. Stockholm: Allduplo Tryckeri AB, Stockholm University.
- Ansaldò, Umberto 2002. Origins and sources of Cocos Malay. NUS-Lund (Sweden) Linguistic Colloquium, National University of Singapore, 29–30 October 2002.
- Ansaldò, Umberto 2003a. Creolization, restructuring and hybridity in the Southeast Asian context. Invited talk at Cornell University, 10 March 2003.
- Ansaldò, Umberto 2003b. Melayu Kokos: A take on pidginization, restructuring and classification. The 7th International Symposium on Malay/Indonesian Linguistics. Nijmegen, The Netherlands, 28–30 June 2003.
- Ansaldò, Umberto 2004a. Contact, typology and the speaker: The essentials of language. *Language Sciences* 26: 485–94.
- Ansaldò, Umberto 2004b. The evolution of Singapore English: Finding the matrix. In Lim (ed.), 127–49.
- Ansaldò, Umberto 2005. Typological admixture in Sri Lanka Malay: The case of Kirinda Java. Ms. University of Amsterdam.
- Ansaldò, Umberto 2007. Review of John McWhorter, 2005, *Defining Creole*. *Journal of Pidgin and Creole Languages* 22.1: 170–6.
- Ansaldò, Umberto 2008. Revisiting Sri Lanka Malay: Genesis and classification. In K. David Harrison, David Rood and Arianne Dwyer (eds.), *A World of Many Voices: Lessons from Documenting Endangered Languages*. Typological Studies in Language. Amsterdam/Philadelphia: John Benjamins, 13–42.
- Ansaldò, Umberto 2009a. Complexity and the age of languages. In Aboh and Smith (eds.).
- Ansaldò, Umberto 2009b. The Asian typology of English: Theoretical and methodological considerations. In Lisa Lim and Nikolas Gisborne (eds.), Special Issue on *The Typology of Asian Englishes*. *English World-Wide* 30(2): 133–48.
- Ansaldò, Umberto and Lisa Lim 2004a. Phonetic absence as syntactic prominence: Grammaticalization in isolating tonal languages. In Olga Fischer, Muriel Norde and Harry Perridon (eds.), *Up and Down the Cline – The Nature of Grammaticalization*. Amsterdam/Philadelphia: John Benjamins, 345–62.
- Ansaldò, Umberto and Lisa Lim 2004b. Kirinda Java and the Malay creoles of Sri Lanka. Curaçao Creole Conference, Curaçao, Netherlands Antilles, 11–15 August 2004.
- Ansaldò, Umberto and Lisa Lim 2006. Globalisation, empowerment and the periphery: The Malays of Sri Lanka. In R. Elangaiyan, R. McKenna Brown, Nicholas D. M. Ostler and Mahendra K. Verma (eds.), *Vital Voices: Endangered Languages and*

- Multilingualism. Proceedings of the FEL X Conference*. Bath: Foundation for Endangered Languages/Mysore: Central Institute of Indian Languages, 39–46.
- Ansaldo, Umberto, Lisa Lim and Salikoko S. Mufwene 2007. The sociolinguistic history of the Peranakans: What it tells us about 'creolization'. In Ansaldo, Matthews and Lim (eds.), 203–26.
- Ansaldo, Umberto and Stephen J. Matthews 1999. The Minnan substrate and creolization in Baba Malay. *Journal of Chinese Linguistics* 27.1: 38–68.
- Ansaldo, Umberto and Stephen J. Matthews 2001. Typical creoles and simple languages: The case of Sinitic. *Linguistic Typology* 5.2/3: 311–26.
- Ansaldo, Umberto and Stephen J. Matthews 2004. The origins of Macanese reduplication. In Geneviève Escure and Armin Schwegler (eds.), *Creoles, Contact and Language Change: Linguistic and Social Implications*. Creole Language Library 27. Amsterdam/Philadelphia: John Benjamins, 1–19.
- Ansaldo Umberto and Stephen Matthews 2007. Deconstructing Creole: The rationale. In Ansaldo, Matthews and Lim (eds.), 1–18.
- Ansaldo, Umberto, Stephen Matthews and Lisa Lim (eds.) 2007. *Deconstructing Creole*. Typological Studies in Language 73. Amsterdam/Philadelphia: John Benjamins.
- Ansaldo, Umberto, Stephen Matthews and Geoff Smith to appear. The Cantonese substrate in China Coast Pidgin. In Claire Lefebvre (ed.), *Substrate Influence in Creole Languages*. Typological Studies in Language. Amsterdam/Philadelphia: John Benjamins.
- Ansaldo, Umberto and Sebastian Nordhoff 2009. Complexity and the age of languages. In Aboh and Smith (eds.).
- Arana-Ward, Marie 1977. A synchronic and diachronic investigation of Macanese: The Portuguese-based creole of Macao. Unpublished MA thesis, University of Hong Kong.
- Arends, Jacques 2001a. Social stratification and network relations in the formation of Sranan. In Norval Smith and Tonjes Veenstra (eds.), *Creolization and Contact*. Amsterdam/Philadelphia: John Benjamins, 291–308.
- Arends, Jacques 2001b. Simple grammars, complex languages. *Linguistic Typology* 5.2/3: 180–2.
- Arends, Jacques, Pieter Muysken and Norval Smith (eds.) 1994. *Pidgins and Creoles: An Introduction*. Amsterdam/Philadelphia: John Benjamins.
- Arvidsson, Stefan 2006. *Aryan Idols: Indo-European Mythology as Ideology and Science*. Chicago/London: The University of Chicago Press.
- Auer, Peter and Li Wei (eds.) 2007. *Multilingualism and Multilingual Communication*. Berlin/New York: Mouton de Gruyter.
- Austin, John L. 1962. *How to Do Things with Words*. London: Oxford University Press.
- Baker, Philip 1990. Off target. *Journal of Pidgin and Creole Languages* 5: 107–19.
- Baker, Philip 1993. Assessing the African contribution to French-based creoles. In Salikoko S. Mufwene (ed.), *Africanisms in Afro-American Language Varieties*. Athens, GA: University of Georgia Press, 123–55.
- Baker, Philip and Chris Corne 1986. Universals, substrata and the Indian Ocean creoles. In Muysken and Smith (eds.), 163–230.
- Baker, Philip and Peter Mühlhäusler 1990. From business to pidgin. *Journal of Asian Pacific Communication* 1.1: 87–115.
- Bakhtin, Mikhail M. 1981. *The Dialogic Imagination*. Ed. and transl. by C. Emerson and M. Holmquist. Austin: University of Texas Press.

- Bakker, Peter 2000. Rapid language change: Creolization, intertwining, convergence. In Colin Renfrew, April McMahon and Larry Trask (eds.), *Time Depth in Historical Linguistics*. Cambridge: McDonald Institute for Archeological Research, 585–620.
- Bakker, Peter and Martin Mous (eds.) 1994. *Mixed Languages: 15 Case Studies in Language Intertwining*. Amsterdam: IFOTT.
- Bao Zhiming 2001. The origins of empty categories in Singapore English. *Journal of Pidgin and Creole Languages* 16.2: 275–319.
- Bao Zhiming 2005. The aspectual system of Singapore English and the systemic sub-stratist explanation. *Journal of Linguistics* 41: 237–67.
- Bao Zhiming and Lye Hui Min 2005. Systemic transfer, topic prominence, and the bare conditional in Singapore English. *Journal of Pidgin and Creole Languages* 20: 269–91.
- Barros, João de and Diogo de Couto 1788. *Da Asia: Nova Edição*. Lisbon: Regia Officina Typographica.
- Batalha, G. N. 1977. *Glossário do Dialecto Macaense: Notas Linguísticas, Etnográficas, Folclóricas*. Coimbra: Instituto de Estudos Românicos.
- Batalha, G. N. 1994. Language of Macau: Past and present. *Review of Culture* 20: 131–56.
- Bauman, Zygmunt 1992. *Intimations of Post-Modernity*. London: Routledge.
- Bausani, Alessandro (ed.) 1972. *L'Indonesia nella Relazione di Viaggio di Antonio Pigafetta*. Roma: Istituto Italiano per il Medio ed Estremo Oriente.
- Bawden, Charles R. 1954. An eighteenth century Chinese source for the Portuguese dialect of Macao. *Zinbun Kagaku Kenkyusyo* (Silver Jubilee vol.): 12–33.
- Baxter, Alan N. 1985. A description of Papia Kristang. Unpublished PhD thesis, Australian National University, Canberra.
- Baxter, Alan N. 1996. Portuguese and Creole Portuguese in the Pacific and Western Pacific Rim. In Stephen A. Wurm, Peter Mühlhäusler and Darrel T. Tryon (eds.), *Atlas of Languages of Intercultural Communication in the Pacific, Asia and the Americas*, vol. II.1. Berlin: Mouton de Gruyter, 299–338.
- Beebe, Leslie M. and Howard Giles 1984. Speech-accommodation theories: A discussion in terms of second language acquisition. *International Journal of the Sociology of Language* 46: 1–32.
- Benjamin, Geoffrey 1988. *Grammar and Polity: The Cultural and Political Background to Standard Malay*. Department of Sociology Working Papers 88, National University of Singapore.
- Bethencourt, Francisco 2005. Low cost empire: Interaction between Portuguese and local societies in Asia. In Ernst van Veen and Leonard Blussé (eds.), *Rivalry and Conflict: European Traders and Asian Trading Networks in the 16th and 17th Centuries*. Leiden: CNWS publications, 108–30.
- Bichsel-Stettler, Anne 1989. Aspects of the Sri Lanka Malay community and its language. Unpublished MA thesis, Bern University, Switzerland.
- Bickerton, Derek 1981. *Roots of Language*. Ann Arbor: Karoma.
- Bickerton, Derek 1984. The Language Bioprogram Hypothesis. *Behavioral and Brain Sciences* 7: 173–221.
- Blake, Barry 1994. *Case*. Cambridge University Press.
- Blommaert, Jan and J. Verschuren 1992. The role of language in European nationalist ideologies. In Bambi B. Schieffelin, Kathryn A. Woolard and Paul V. Kroskrity (eds.), Special Issue on *Language Ideologies*. *Pragmatics* 2.3: 235–453.

- Bloomfield, Leonard 1933. *Language*. New York: Henry Holt & Co.
- Bloomfield, Leonard 1944. Secondary and tertiary responses to language. *Language* **20**: 44–55.
- Boas, Franz 1911. Introduction. In Franz Boas (ed.), *Handbook of North American Indian Languages. Part 1. Bulletin of the Bureau of American Ethnology*, vol. 40. Washington DC: Government Printing Office, 1–83.
- Bodman, Nicholas C. 1955. *Spoken Amoy Hokkien*, vol. I. Kuala Lumpur: Charles Grenier.
- Bolton, Kingsley 2003. *Chinese Englishes: A Sociolinguistic History*. Cambridge University Press.
- Bopp, Franz 1857. *Vergleichende Grammatik des Sanskrit, Zend, Griechischen, Lateinischen, Litauischen, Gotischen und Deutschen*. 2nd edn. Berlin: Fred Dümmler's Verlagsbuchhandlung.
- Bourdieu, Pierre 1977. *Outline of a Theory of Practice*. Cambridge University Press.
- Boxer, Charles R. (ed.) 1953. *South China in the Sixteenth Century*. London: Hakluyt Society.
- Boxer, Charles R. 1965. *Portuguese Society in the Tropics: The Municipal Councils of Goa, Macao, Bahia, and Luanda, 1501–1800*. Madison: University of Wisconsin Press.
- Boxer, Charles R. 1973. *The Portuguese Seaborne Empire*. Hammonds Worth: Pelican.
- Buchholtz, Mary and Kira Hall 2007. Language and identity. In Alessandro Duranti (ed.), *A Companion to Linguistic Anthropology*. Malden, MA: Blackwell, 369–94.
- Bunce, Pauline 1988. *The Cocos (Keeling) Islands: Australian Atolls in the Indian Ocean*. Singapore: The Jacaranda Press.
- Burah, Tuan Arfin 2006. *Saga of the Exiled Royal Javanese Unearthed*. Dehiwala, Sri Lanka.
- Bybee, Joan 1985. *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam/Philadelphia: John Benjamins.
- Bybee, Joan 1998. A functionalist approach to grammar and its evolution. *Evolution of Communication* **2**: 249–78.
- Bybee, Joan 2001. *Phonology and Language Use*. Cambridge University Press.
- Bybee, Joan 2006. From usage to grammar: The mind's response to repetition. *Language* **82.4**: 711–33.
- Bybee, Joan 2007. *Frequency of Use and the Organization of Language*. Oxford University Press.
- Bybee, Joan and Paul Hopper (eds.) 2001. *Frequency and the Emergence of Linguistic Structure*. Amsterdam/Philadelphia: John Benjamins.
- Bybee, Joan and Sandra A. Thompson 1997. Three frequency effects in syntax. *Berkeley Linguistic Society* **23**: 378–88.
- Cabral, Joao de Pina 1994. The ethnic composition of Macao. *Review of Culture* **20**: 229–39.
- Campbell, Lyle 2002. *Historical Linguistics: An Introduction*. Cambridge: MIT Press.
- Cavalli-Sforza, Luigi 2000. *Genes, People and Languages*. New York: North Point Press.
- Caws, Peter 2000 *Structuralism: A Philosophy for the Human Sciences*. New York: Humanity Books.
- Chaudenson, Robert 1979. *Les Creoles Français*. Evreux: Nathan.
- Chaudenson, Robert 2001. (Revised in collaboration with S. S. Mufwene.) *Creolization of Language and Culture*. London: Routledge.

- Chen, Ta 1939. *Emigrant Communities in South China: A Study of Overseas Migration and its Influence on Standards of Living and Social Change*. Shanghai: Kelly & Walsh Ltd.
- Chomsky, Noam 1959. Review of Verbal Behaviour by B. F. Skinner. *Language* 35: 26–58.
- Chomsky, Noam 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Cifoletti, Guido 2004. *La Lingua Franca Barbaresca*. Roma: Il Calamo.
- Clements, Clancy 1992. Foreigner Talk and the origins of pidgin Portuguese. *Journal of Pidgin and Creole Languages* 7.1: 75–92.
- Clements, Clancy 2003. An amorphous morphological account of reduplication processes in three Indo-Portuguese Creoles. In Silvia Kouwenberg (ed.), *Reduplication in Pidgins and Creoles*. London: Battlebridge Press, 193–202.
- Clements, Clancy 2009. *The Linguistic Legacy of Spanish and Portuguese: Colonial Expansion and Language Change*. Cambridge University Press.
- Coates, Austin 1978. *A Macau Narrative*. Oxford University Press.
- Collins, James T. 1984. Malaysian and Bazaar Malay: Polarity, continuity and communication. In A. H. Omar (ed.), *National Language and Communication in Multilingual Societies*. Kuala Lumpur: Dewan Bahasa dan Pustaka, 151–74.
- Collins, James T. 1995. *Bibliografi Dialek Melayu di Pulau Jawa, Bali dan Sri Lanka*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Collins, James T. 1996. *Malay, World Language: A Short History*. Malaysia: Dewan Bahasa dan Pustaka.
- Collins, James T. 1998. *Malay, World Language: A Short History*. 2nd edn. Malaysia: Dewan Bahasa dan Pustaka.
- Collins, James T. 2001. Dutch scholarship and the study of Malay dialects: A sketch of pre-twentieth century contributions. Seminar on the Regional Malays of Indonesia, SIL Indonesia, Yogyakarta, January 2001.
- Corne, Chris 1999. *From French to Creole: The Development of New Vernaculars in the French Colonial World*. London: University of Westminster Press.
- Coulmas, Florian 1988. *With Forked Tongues. What are National Languages Good for?* Ann Arbor, MI: Karoma.
- Coulmas, Florian 1989. *The Writing Systems of the World*. Oxford: Blackwell.
- Couto, Diego do 1645. *Da Asia*. Decades. Lisbon: Regia Officina Typografica 1778–88. [Reprinted Lisbon, 1974].
- Crawford, John 1856. *A Descriptive Dictionary of the Indian Island and Adjacent Countries*. London: Bradbury & Evans.
- Croft, William 1990. *Typology and Universals*. Cambridge University Press.
- Croft, William 1995. Autonomy and functionalist linguistics. *Language* 71: 490–532.
- Croft, William 2000. *Explaining Language Change: An Evolutionary Approach*. Edinburgh: Longman.
- Croft, William 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford University Press.
- Croft, William 2003. Social evolution and language change. Ms. University of Manchester.
- Croft, William 2006a. Evolutionary models and functional-typological theories of language change. In Ans v. Kamenade and Bettelou Los (eds.), *Handbook of the History of English*. Oxford: Blackwell, 68–91.

- Croft, William 2006b. The relevance of an evolutionary model. In Ole N. Thomsen (ed.), *Competing Models of Linguistic Change: Evolution and Beyond*. Amsterdam/Philadelphia: John Benjamins, 91–132.
- Dahl, Östen 2004. *The Growth and Maintenance of Linguistic Complexity*. Amsterdam/Philadelphia: John Benjamins.
- Dalgado, Sebastiao R. 1917. Dialecto Indo-Português de Negapatão. *Revista Lusitana* 20: 40–53.
- Dampier, William 1699. *Voyages and Discoveries*. [ed. by Clennell Wilkinson 1931.] London: Argonaut Press.
- Darwin, Charles 1842. *The Structure and Distribution of Coral Reefs. Being the First Part of the Geology of the Voyage of the Beagle, under the Command of Capt. Fitzroy, R. N. during the Years 1832 to 1836*. London: Smith, Elder & Co.
- Darwin, Charles 1859. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray.
- Dawkins, Richard 1976. *The Selfish Gene*. New York: Oxford University Press.
- DeGraff, Michel 1999. Creolization, language change and language acquisition: An epilogue. In Michel DeGraff (ed.), *Language Creation and Language Change: Creolization, Diachrony and Development*. Cambridge, MA: MIT Press, 473–543.
- DeGraff, Michel 2001. On the origins of creoles: A Cartesian critique of Neo-Darwinian linguistics. *Linguistic Typology* 5: 213–30.
- DeGraff, Michel 2002. Relexification: A re-evaluation. *Anthropological Linguistics* 44.4: 321–414.
- DeGraff, Michel 2003. Against Creole exceptionalism. Discussion note. *Language* 79: 391–410.
- DeGraff, Michel 2005. Linguists' most dangerous myth: The fallacy of Creole exceptionalism. *Language in Society* 34: 533–91.
- Den Besten, Hans 2007. Relexification and pidgin development: The case of Cape Dutch Pidgin. In Ansaldo, Matthews and Lim (eds.), 141–64.
- Den Besten, Hans, Pieter Muysken and Norval Smith 1994. Theories focusing on the European input. In Arends, Muysken and Smith (eds.), 87–98.
- Dixon, R. M. W. 1997. *The Rise and Fall of Languages*. Cambridge University Press.
- Djité, Paulin G. 2006. Shifts in linguistic identities in a global world. *Language Problems and Language Planning* 30.1: 1–20.
- Dreyfuss, Gail Raimi and Djoejana Oka 1979. Chinese Indonesian: A new kind of language hybrid? *Papers in Pidgin and Creole Linguistics No. 2*. Pacific Linguistics Series A-57: 247–74.
- Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68: 81–138.
- Dryer, Matthew S. 2003. Word order in Sino-Tibetan languages from a typological and geographical perspective. In Graham Thurgood and Randy LaPolla (eds.), *Sino-Tibetan Languages*. Richmond: Curzon Press, 43–55.
- Dumont, Louis 1970. *Homo Hierarchicus: The Caste System and its Implications*. Transl. by M. Sainsbury. London: Weidenfeld and Nicolson.
- Duranti, Alessandro 1997. *Linguistic Anthropology*. Cambridge University Press.
- Duranti, Alessandro (ed.) 2001. *Linguistic Anthropology: A Reader*. United Kingdom: Blackwell.



- Durham, William H. 1991. *Coevolution, Genes, Culture and Human Diversity*. Palo Alto: Stanford University Press.
- Durie, Mark and Malcolm Ross (eds.) 1996. *The Comparative Method Reviewed: Regularity and Irregularity in Language Change*. New York: Oxford University Press.
- Eckert, Penelope 1980. Diglossia: Separate and unequal. *Linguistics* **18**: 156–64.
- Eckert, Penelope 1989. The whole woman: Sex and gender differences in variation. *Language Variation and Change* **1**: 245–67.
- Eckert, Penelope and Sally McConnell-Ginet 1992. Think practically and look locally: Language and gender as community-based practice. *Annual Review of Anthropology* **21**: 461–90.
- Edelman, Gerald M. 1992. *Bright Air, Brilliant Fire: On the Matter of the Mind*. New York: Basic Books.
- Edwards, John 1994. *Multilingualism*. London: Routledge.
- Emeneau, Murray B. 1980. *Language and Linguistic Area*. Stanford University Press.
- Enfield, Nicholas J. 2003. *Linguistic Epidemiology: Semantics and Grammar of Language Contact in Mainland Southeast Asia*. London: Routledge.
- Enfield, Nicholas J. 2008. Transmission biases in linguistic epidemiology. *Journal of Language Contact – THEMA* **2**: 299–310. [www.jlc-journal.org](http://www.jlc-journal.org).
- Errington, John J. 1992. On the ideology of Indonesian language development: The state of a language of state. In Bambi B. Schieffelin, Kathryn A. Woolard and Paul V. Kroskrity, (eds.), Special Issue on Language Ideologies. *Pragmatics* **2.3**: 235–453.
- Escure, Genevieve to appear. Is serialization simple? Evidence from Chinese Pidgin English. In Tomas Klein and Nicolas Faraclas (eds.), *Simplification in Pidgins and Creoles*. Amsterdam/Philadelphia: John Benjamins.
- Fairbank, John K. 1953. *Trade and Diplomacy on the China Coast: The Opening of the Treaty Ports, 1842–1854*. Stanford University Press.
- Fairbank, John K. and Merle Goldman 1998. *China: A New History*. 2nd enlarged edn. Cambridge, MA/London, UK: The Belknap Press of Harvard University Press.
- Faraclas, Nicholas, Micah Corum, Rhoda Arrindell and Jean O. Pierre 2007. Sociétés de cohabitation and the similarities between the English-lexifier Creoles of the Atlantic and the Pacific: The case for diffusion from the Afro-Atlantic to the Pacific. 24th Meeting of the Society for Pidgin and Creole Linguistics, Amsterdam, 18–20 June 2007.
- Faraclas, Nicholas, Don E. Walicek, Mervyn C. Alleyne, Wilfredo Geigel and Luis Ortiz 2007. The complexity that really matters: The role of political economy in creole genesis. In Ansaldo, Matthews and Lim (eds.), 227–64.
- Fels, Susan (ed.) 1999. *Before the Wind: The Memoir of an American Sea Captain, 1808–1833, by Charles Tyng*. New York: Penguin.
- Ferguson, Charles A. 1971. Absence of copula and the notion of simplicity: A study of normal speech, baby talk, foreigner talk and pidgins. In Hymes (ed.), 141–50.
- Ferguson, Charles A. 1975. Towards a characterization of Foreigner Talk. *Anthropological Linguistics* **17**: 1–14.
- Ferguson, Charles A. 1982. Foreword. In Braj Kachru (ed.), *The Other Tongue: English Across Cultures*. Oxford: Pergamon, vii–xi.
- Fernandes, Miguel S. and Alan Baxter 2001. *Maquista Chapado: Vocabulário e Expressões de Crioulo Português de Macau*. Macau: Instituto Internacional de Macau.



- Ferreira, José dos Santos 1978. *Papiá Cristám de Macau*. Macau: Serviços de Turismo.
- Ferreira, José dos Santos 1983. *Poéma di Macau*. Macau: Leal Senado.
- Ferreira, José dos Santos 1996. *Papiaçám di Macau*. Macau: Fundação Macau.
- Feyerabend, Paul 1975. *Against Method: Outline of an Anarchistic Theory of Knowledge*. London/New York: New Left Books.
- Field, Fred 2002. *Linguistic Borrowing in Bilingual Contexts*. Studies in Language 62. Amsterdam/Philadelphia: John Benjamins.
- Fishman, Joshua A. 1964. Language maintenance and language shift as a field of enquiry: A definition of the field and suggestions for its further development. *Linguistics* 9: 32–70.
- Fishman, Joshua A. 1989. *Language and Ethnicity in Minority Sociolinguistic Perspective*. Clevedon: Multilingual Matters.
- Fishman, Joshua A. 1991. *Reversing Language Shift: Theoretical and Empirical Foundations of Assistance to Threatened Languages*. Clevedon: Multilingual Matters.
- Fitch, Tecumseh W. 2007. An invisible hand. *Nature* 449.11: 665–7.
- Foley, William 1986. *The Papuan Languages of New Guinea*. Cambridge University Press.
- Forbes, Henry O. 1885. *A Naturalist's Wanderings in the Eastern Archipelago: A Narrative of Travel and Exploration from 1878 to 1883*. New York: Harper & Brothers.
- Foster, William (ed.) 1944. *The Voyage of Sir Henry Middleton to the Moluccas 1604–1606*. London: Hakluyt Society.
- Foucault, Michel 1972. *The Archeology of Knowledge*. New York: Harper and Row.
- Foucault, Michel 1980. *Power/Knowledge: Selected Interviews and Other Writings, 1972–1977*. New York: Pantheon.
- Fox, James J. 2005. Ritual languages, special registers and speech decorum in Austronesian languages. In Karl Alexander Adelaar and Nikolaus P. Himmelmann (eds.), *The Austronesian Languages of Asian and Madagascar*. London and New York: Routledge, 87–109.
- Gair, James and John Paolillo 1997. *Sinhala*. Munich: Lincom Europa.
- Galvão, Antonzo 1544 [1971]. *A Treatise on the Moluccas (c. 1544). Probably the Preliminary Version of Antonio Galvão's Lost História das Molucas*. Transl. by H. Jacobs. Rome: Jesuit Historical Institute.
- Garrett, Paul 2000. 'High' Kwéyòl: The emergence of a formal Creole register in St. Lucia. In John H. McWhorter (ed.), *Language Change and Language Contact in Pidgins and Creoles*. Amsterdam/Philadelphia: John Benjamins, 63–101.
- Garrett, Paul B. 2006. Contact languages as "endangered" languages: What is there to lose? *Journal of Pidgin and Creole Languages* 21.1: 175–90.
- Garrett, Paul and Patricia Baquedano-López 2002. Language socialization: Reproduction and continuity, transformation and change. *Annual Review of Anthropology* 31: 339–61.
- Geertz, Clifford 1973. *The Interpretation of Cultures*. New York: Basic Books.
- Geertz, Clifford 1983. *Local Knowledge. Further Essays in Interpretative Anthropology*. New York: Basic Books.
- Gelber, Harry G. 2007. *The Dragon and the Foreign Devils: China and the World, 1100 BC to the Present*. London: Bloomsbury.
- Gibson-Hill, Carl Alexander 1947. Notes on the Cocos-Keeling Islands. *Journal of the Malayan Branch of the Royal Asiatic Society* XX.2: 140–202.

- Gil, David 2000. Syntactic categories, cross-linguistic variation and Universal Grammar. In Petra M. Vogel and Bernard Comrie (eds.), *Approaches to the Typology of Word Classes*. Berlin/New York: Mouton, 173–216.
- Gil, David 2001. Escaping Eurocentrism: Fieldwork as a process of unlearning. In Paul Newman and Martha Ratliff (eds.), *Linguistic Fieldwork*. Cambridge University Press, 102–32.
- Giles, Howard 1980. Accommodation Theory: Some new directions. *York Papers in Linguistics* 9.
- Givón, Talmy 1979a. *On Understanding Grammar*. New York: Academic Press.
- Givón, Talmy 1979b. Prolegomena to any sane creology. In Ian F. Hancock (ed.), *Readings in Creole Studies*. Ghent: Story Scientia, 3–35.
- Givón, Talmy 1991. Markedness in grammar: Distributional, communicative and cognitive correlates of syntactic structure. *Studies in Language* 15.2: 335–70.
- Givón, Talmy 1995. *Functionalism and Grammar*. Amsterdam/Philadelphia: John Benjamins.
- Glynn, John H. (ed.) 1998. *Indonesian Heritage: Language and Literature*. Singapore: Archipelago Press.
- Goddard, Cliff 2005. *The Languages of East and Southeast Asia*. Oxford University Press.
- Goldinger, Stephen D. 1996. Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition* 22.1: 1166–83.
- Gomes, Luis G. 1950 *Monografia de Macau*. Macao: Imprensa Nacional.
- Goodman, Morris 1987. Pidgins origins reconsidered. *Journal of Pidgin and Creole Languages* 2: 149–62.
- Gordon, Raymond G. Jr. (ed.) 2005. *Ethnologue: Languages of the World*. 15th edn. Dallas, TX: SIL International. <http://ethnologue.com>. Accessed 11 March 2007.
- Graaf, Hermanus J. de and Theodore G. Pigeaud 1974. *De Eerste Moslims Vorstendommen op Java. Studien over de Staatkundige Geschiedens van de 15de en 16de Eeuw*. The Hague: Nijhoff KITLV.
- Grant, Anthony 2007. Admixture and after: The Chamic languages and the creole prototype. In Ansaldo, Matthews and Lim (eds.), 109–40.
- Greenberg, Joseph 1966. Language universals, with special reference to feature hierarchies. *Janua Linguarum*, Series Minor 59. The Hague: Mouton.
- Grijns, C.D. 1991. *Jakarta Malay: A Multidimensional Approach to Spatial Variation*. Leiden: KITLV Press.
- Gumperz, John J. 1958. Dialect differences and social stratification in a North Indian village. *American Anthropologist* 60: 668–82.
- Gumperz, John J. 1964. Linguistic and social interaction in two communities. *American Anthropologist* 66.6: 137–53.
- Gumperz, John J. 1968. The speech community. *International Encyclopedia of the Social Sciences*. New York: Macmillan, 381–86.
- Gumperz, John J. and Robert Wilson 1971. Convergence and creolization: A case from the Indo-Aryan/Dravidian border in India. In Hymes (ed.), 151–67.
- Gunn, Geoffrey C. 1996. *Encountering Macau: A Portuguese City-State on the Periphery of China, 1557–1999*. Westview Press: Harper Collins.
- Gupta, Anthea Fraser 1994. *The Step-Tongue: Children's English in Singapore*. Clevedon/Philadelphia/Adelaide: Multilingual Matters Ltd.

- Haan, Frederik de 1922. *Oud Batavia. Gedenkboek*. Batavia: Kolff.
- Hagège, Claude 1993. *The Language Builder: An Essay on the Human Signature in Linguistic Morphogenesis*. Amsterdam/Philadelphia: John Benjamins.
- Haiman, John 1994. Ritualization and the development of language. In William Pagliuca (ed.), *Perspectives on Grammaticalization*. Amsterdam/Philadelphia: John Benjamins, 3–28.
- Hall, Kenneth R. 1985. *Maritime Trade and State Development in Early Southeast Asia*. Honolulu: University of Hawaii Press.
- Hall, Robert A. 1944. Chinese pidgin grammar and texts. *Journal of the American Oriental Society* 64.3: 95–115.
- Hancock, Ian F. 1975. Melacca Creole Portuguese: Asian, African or European? *Anthropological Linguistics* 17.1: 211–36.
- Harris, Roy 1980. *The Language Makers*. London: Duckworth.
- Harris, Roy 2004. *The Linguistics of History*. Edinburgh University Press.
- Haspelmath, Martin 1998. How young is Standard Average European? *Language Sciences* 20.3: 271–87.
- Haspelmath, Martin 2006. Against markedness (and what to replace it with). *Journal of Linguistics* 42.1: 25–70.
- Haspelmath, Martin 2008. Terminology of case. In Andrej Malchukov and Andrew Spencer (eds.), *Handbook of Case*. Oxford University Press.
- Haugen, Einar 1972. *The Ecology of Language: Essays by Einar Haugen*. (Selected and introduced by Anwar S. Dil.) Stanford University Press.
- Hauser, Marc D., Noam Chomsky and W. Tecumseh Fitch 2002. The faculty of language: What is it, who has it and how did it evolve? *Science* 298: 1569–79.
- Hawkins, John A. 1994. *A Performance Theory of Order and Constituency*. Cambridge University Press.
- Heine, Bernd and Tania Kuteva 2005. *Language Contact and Grammatical Change*. Cambridge University Press.
- Heugh, Kathleen and Christopher Stroud 2004. Language rights and linguistic citizenship. In Jane Freeland and D. Patrick (eds.), *Language Rights and Language Survival: A Sociolinguistic Exploration*. Manchester: St Jerome Publishing, 191–218.
- Hill, Jane H. and Kenneth C. Hill 1986. *Speaking Mexicano: Dynamics of Syncretic Language in Central Mexico*. Tucson: University of Arizona Press.
- Himmelman, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. In Karl Alexander Adelaar and Nikolaus P. Himmelman (eds.), *The Austronesian Languages of Asian and Madagascar*. London/New York: Routledge, 110–81.
- Hjelmslev, Louis 1938. Relation de parenté des langues creoles. *Revue des Études Indo-Européennes* 2: 271–86.
- Hoadley, Mason C. 1983. Slavery, bondage and dependence in pre-colonial Java. The Cirebon-Priangan region, 1700. In Reid (ed.), 90–117.
- Holm, John 1989. *Pidgins and Creoles: A Reference Survey*, vol. II. Cambridge University Press.
- Hopper, Paul 1987. Emergent grammar. *Berkeley Linguistic Society* 13: 139–57.
- Hopper, Paul and Elisabeth Traugott 1993. *Grammaticalization*. Cambridge University Press.
- Hull, David 1988. *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. University of Chicago Press.

- Hull, David 2001. *Science and Selection: Essays on Biological Evolution and the Philosophy of Science*. Cambridge University Press.
- Hull, David, Rodney E. Langman and Sigrid S. Glenn 2001. A general account of selection. In Hull, 49–96.
- Hunter, William C. 1885. *Bits of Old China*. London: Kegan, Paul, Trench and Co.
- Hussainmiya, B.A. 1987. *Lost Cousins: The Malays of Sri Lanka*. Kuala Lumpur: Universiti Kebangsaan Malaysia.
- Hussainmiya, B.A. 1990. *Orang Rejimen: The Malays of the Ceylon Rifle Regiment*. Kuala Lumpur: Universiti Kebangsaan Malaysia.
- Hutton, Christopher 1998. *Linguistics and the Third Reich: Mother-Tongue Fascism, Racism and the Science of Language*. London: Routledge.
- Hymes, Dell (ed.) 1971. *Pidginization and Creolization of Languages*. Cambridge University Press.
- Hymes, Dell 1972. On communicative competence. In Janet B. Pride and John Holmes (eds.), *Sociolinguistics*. Harmondsworth: Penguin, 269–93.
- Ingold, Tim 2000. *The Perception of the Environment: Essays in Livelihood, Dwelling and Skill*. London/New York: Routledge.
- Jakobson, Roman 1963. Implications of language universals for linguistics. In Joseph Greenberg (ed.), *Universals of Language*. MIT Press, 263–78.
- Janda, Richard and Brian Joseph 2003. On language, change and language change – or of history, linguistics and historical linguistics. In Brian Joseph and Richard Janda (eds.), *Handbook of Historical Linguistics*. Oxford: Blackwell, 3–180.
- Johnson, Allen W. and Timothy Earle 2000. *The Evolution of Human Societies: From Foraging Group to Agrarian State*. Stanford University Press.
- Jourdan, Christine 1991. Pidgins and creoles: The blurring of categories. *Annual Review of Anthropology* 20: 187–209.
- Kachru, Braj B. 1996. The paradigms of marginality. *World Englishes* 15: 241–55.
- Kapuściński, Ryszard 2004. *Travels with Herodotus*. New York: Alfred A. Knopf.
- Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
- Keller, Rudi 1994. *On Language Change: The Invisible Hand in Language*. London: Routledge.
- Khin Khin Aye 2005. A grammar of Singapore Bazaar Malay. Unpublished PhD thesis, National University of Singapore.
- Knappert, Jan 1985. East Africa and the Indian Ocean. In J.C. Stone (ed.), *Africa and the Sea: Proceedings of a Colloquium at the University of Aberdeen, March 1984*. University of Aberdeen, 125.
- Kobata, Atsushi and Mitsugu Matsuda 1969. *Ryukyuan Relations with Korea and South Sea Countries*. Kyoto: Atsushi Kobata.
- Koptjevskaja-Tamm, Maria to appear. Linguistic typology and language contact. In Jae Jung Song (ed.), *The Oxford Handbook of Typology*. Oxford University Press.
- Kroskrity, Paul V. (ed.) 2000. *Regimes of Language: Ideologies, Politics and Identities*. Santa Fe: School of American Research Press.
- Kuhn, Thomas S. 1970. *The Structure of Scientific Revolutions*. 2nd edn. University of Chicago Press.
- Kulick, Don and Bambi Schieffelin 2007. Language socialization. In Duranti (ed.), 349–68.
- Kusters, Christian W. 1996. *Linguistic Complexity: The Influence of Social Change on Verbal Inflection*. PhD thesis, The Netherlands: LOT.

- Lakatos, Imre 1970. Falsification and the methodology of scientific research programs. In Imre Lakatos and John Worrall (eds.), *Criticism and the Growth of Knowledge*. Cambridge University Press, 91–195.
- Langacker, Ronald 1987. *Foundations of Cognitive Grammar*, vol. I. Stanford University Press.
- Langacker, Ronald 2000. A dynamic usage-based model. In Michael Barlow and Suzanne Kemmer (eds.), *Usage-Based Models of Language*. Stanford: Center for the Study of Language and Information, 1–63.
- Lapsley, Alistair D. 1983. Cocos Malay syntax. Unpublished MA thesis, Monash University, Melbourne.
- Lass, Roger 1997. *Historical Linguistics and Language Change*. Cambridge University Press.
- Leather, Jonathan and Jet van Dam (eds.) 2003. *Ecology of Language Acquisition*. The Netherlands: Kluwer.
- Lee, Mae-En Gwyneth Adele 2000. A descriptive grammar of spoken Peranakan. Unpublished BA Honours thesis, National University of Singapore.
- Lefebvre, Claire 1998. *Creole Genesis and the Acquisition of Grammar: The Case of Haitian Creole*. Cambridge University Press.
- Lefebvre, Claire 2001. Relexification in creole genesis and its effects on the development of the creole. In Norval Smith and Tonjes Veenstra (eds.), *Creolization and Contact*. Amsterdam/Philadelphia: John Benjamins, 9–42.
- Leland, Charles 1892. *Pidgin English Sing-Song*. London: Kegan Paul.
- Le Page, Robert and Andrée Tabouret-Keller 1985. *Acts of Identity: Creole-Based Approaches to Language and Ethnicity*. Cambridge University Press.
- Leur, Jacob C. van 1955. *Indonesian Trade and Society*. The Hague: W. van Houve.
- Levathes, Louise 1994. *When China Ruled the Seas: The Treasure Fleet of the Dragon Throne, 1405–33*. New York: Oxford University Press.
- Li, Michelle, Stephen Matthews and Geoff P. Smith 2005. Pidgin English texts from the Chinese English Instructor. In Geoff P. Smith and Stephen Matthews (eds.), Special Issue on *Chinese Pidgin English: Texts and Contexts*. *Hong Kong Journal of Applied Linguistics* 10.1: 79–168.
- Lieberman, Erez, Jean-Baptiste Michel, Joe Jackson, Tina Tang and Martin A. Nowak 2007. Quantifying the evolutionary dynamics of language. *Nature* 449.11: 713–6.
- Lieberman, Philip 2006. *Toward an Evolutionary Biology of Language*. Cambridge, MA/London, England: Belknap Harvard University Press.
- Lightfoot, David 1999. *The Development of Language: Acquisition, Change and Evolution*. Oxford: Blackwell.
- Lim, Boon Keng 1917. The Chinese in Malaya. In W. Feldwisk (ed.), *Present Days' Impressions of the Far East and Prominent and Progressive Chinese at Home and Abroad: The History, People, Commerce, Industries and Resources of China, Hong Kong, Indo-China, Malaya and Netherlands India*. London: Globe Encyclopedia Co., 875–82.
- Lim, Lisa (ed.) 2004. *Singapore English: A Grammatical Description*. Varieties of English Around the World G33. Amsterdam/Philadelphia: John Benjamins.
- Lim, Lisa 2007. Mergers and acquisitions: On the ages and origins of Singapore English particles. *World Englishes* 26.4: 446–73.
- Lim, Lisa 2008. Dynamic linguistic ecologies of Asian Englishes. *Asian Englishes* 11.1: 52–5.

- Lim, Lisa 2009a. Revisiting English prosody: (Some) New Englishes as tone languages. In Lisa Lim and Nikolas Gisborne (eds.), Special Issue on *The Typology of Asian Englishes*. *English World-Wide* 30(2): 218–39.
- Lim, Lisa 2009b. Peranakan English. In Daniel Schreier, Peter Trudgill, Edgar Schneider and Jeff Williams (eds.), *Lesser Known Varieties of English*. Cambridge University Press.
- Lim, Lisa forthcoming. Migrants and ‘mother tongues’: Extralinguistic forces in the ecology of English in Singapore. In Lisa Lim, Anne Pakir and Lionel Wee (eds.), *English in Singapore: World Language and Lingua Franca*. Asian Englishes Today. Hong Kong University Press.
- Lim, Lisa and Umberto Ansaldo 2006a. Keeping Kirinda vital: The endangerment-empowerment dilemma in the documentation of Sri Lanka Malay. *Amsterdam Center for Language and Communication Working Papers* 1: 51–66.
- Lim, Lisa and Umberto Ansaldo 2006b. The mechanisms of contact must be complex: Singapore English particles as a case of categorial fusion. The 12th Annual Conference of the International Association for World Englishes (IAWE). World Englishes in World Contexts: Theory and application. Nagoya, Japan, 7–9 October 2006.
- Lim, Lisa and Umberto Ansaldo 2007. Identity alignment in the multilingual space: The Malays of Sri Lanka. In Eric A. Anchimbe (ed.), *Linguistic Identity in Multilingual Postcolonial Spaces*. Cambridge Scholars Press, 218–43.
- Lim, Lisa and Joseph A. Foley 2004. English in Singapore and Singapore English: Background and methodology. In Lim (ed.), 1–18.
- Lim, Sonny 1988. Baba Malay: The language of the ‘Straits-born’ Chinese. In Heinrich Steinhauer (ed.), *Papers in Western Austronesian Linguistics*, no. 3. Pacific Linguistics A-78. Canberra: The Australian National University, 1–61.
- Lindblom, Björn, Peter McNeilage and Michael Studdert-Kennedy 1984. Self-organizing process and the explanation of phonological universals. In Brian Butterworth, Bernard Comrie and Östen Dahl (eds.), *Explanations for Language Universals*. New York: Mouton, 181–203.
- Lombard, Denys 1981. Questions on the contact between European companies and Asian societies. In L. Blussé and F. Gaastra (eds.), *Companies and Trade*. Leiden University Press. 184–85.
- Ma Huan 1433 [1970]. Ying yai sheng lan [Overall survey of the ocean’s shores]. In J. V. G. Mills (ed.), *The Overall Survey of the Ocean’s Shores*. Cambridge: Hakluyt Society.
- McPherson, Kenneth 1993. *The Indian Ocean: A History of People and the Sea*. Oxford University Press.
- McWhorter, John 1998. Identifying the creole prototype: Vindicating a typological class. *Language* 74: 788–818.
- McWhorter, John 2004. Saramaccan and Haitian as young grammars: The pitfalls of syntactocentrism in Creole genesis research. *Journal of Pidgin and Creole Languages* 19.1: 77–138.
- McWhorter, John 2005. *Defining Creole*. Oxford University Press.
- McWhorter, John 2007. *Language Interrupted: Signs of Non-Native Acquisition in Standard Language Grammars*. Oxford University Press.
- Manguin, Pierre-Yves 1983. Manpower and labour in early sixteenth century Melacca. In Reid (ed.), 210–15.



- Marsden, William 1812 [reprint 1984]. *A Grammar and Dictionary of the Malayan Language*. Singapore: Oxford University Press.
- Martino, Elena 2003. Chinese Pidgin English: Genesi ed evoluzione (Chinese Pidgin English: Genesis and evolution). 'Tesi di Laurea', Roma, La Sapienza.
- Masica, Colin 1976. *Defining a Linguistic Area: South Asia*. University of Chicago Press.
- Matras, Yaron 2000. Fusion and the cognitive basis for bilingual discourse markers. *International Journal of Bilingualism* 4.4: 505–28.
- Matras, Yaron 2003. Mixed languages: Re-examining the structural prototype. In Matras and Bakker (eds.), 151–76.
- Matras, Yaron and Peter Bakker (eds.) 2003. *The Mixed Language Debate: Theoretical and Empirical Advances*. Berlin/New York: Mouton de Gruyter.
- Matthews, Stephen and Geoff Smith 2005. Substrate influence as reflected in Chinese and English sources for China Coast Pidgin. Creole Language Structure: Between Superstrate and Substrate, MPI Leipzig, June 2005.
- Matthews, Stephen and Virginia Yip 1994. *Cantonese: A Comprehensive Grammar*. London: Routledge.
- Meillet, Antoine 1948. *Linguistique Historique et Linguistique Générale*. Collection linguistique publiée par la Société de Linguistique de Paris VIII. Paris: Edouard Champion.
- Milroy, James 1999. Roger Lass, historical linguistics and language change. *Diachronica* 16.1: 179–91.
- Milroy, Lesley 1987. *Language and Social Networks*. 2nd edn. Oxford University Press.
- Milroy, Lesley and James Milroy 1992. Social network and social class: Towards an integrated sociolinguistic model. *Language in Society* 21: 1–26.
- Minde, Don van 1997. *Malayu Ambong: Phonology, Morphology, Syntax*. Leiden: CNWS.
- Mintz, Malcolm 1994. *A Student's Grammar of Malay and Indonesian*. Singapore: SNP Publishing.
- Montesquieu, de Bordenave 1949. *The Spirit of the Laws*. Transl. by T. Nugent. New York: Hafner.
- Morse, Hosea B. 1926. *The Chronicles of the East India Company Trading to China, 1635–1834*. 5 vols. Oxford: Clarendon Press.
- Mufwene, Salikoko S. 1990. Creoles and universal grammar. *Linguistics* 28.4: 783–807.
- Mufwene, Salikoko S. 1996a. Creole genesis: A population genetics perspective. In Pauline Christie (ed.), *Caribbean Language Issues Old and New*. Kingston, Jamaica: University of the West Indies Press, 168–209.
- Mufwene, Salikoko S. 1996b. The founder principle in creole genesis. *Diachronica* 13: 83–134.
- Mufwene, Salikoko S. 1998. What research on creole genesis can contribute to historical linguistics. In Monica S. Schmid, Jennifer R. Austin, and Dieter Stein (eds.), *Historical Linguistics 1997*. Amsterdam/Philadelphia: John Benjamins, 315–38.
- Mufwene, Salikoko S. 2000. Creolization is a social, not a structural, process. In Ingrid Neumann-Holzschuh and Edgar W. Schneider (eds.), *Degrees of Restructuring in Creole Languages*. Amsterdam/Philadelphia: John Benjamins, 65–84.
- Mufwene, Salikoko S. 2001. *The Ecology of Language Evolution*. Cambridge University Press.

- Mufwene, Salikoko S. 2002a. Competition and selection in language evolution. *Selections* 3.1: 45–56.
- Mufwene, Salikoko S. 2002b. Colonization, globalization, and the future of languages in the twenty-first century. *MOST Journal on Multicultural Societies* 4.2: 162–93.
- Mufwene, Salikoko S. 2003. Genetic linguistics and genetic creolistics: A response to Sarah G. Thomason's "Creoles and genetic relationships". *Journal of Pidgin and Creole Languages* 18.2: 273–88.
- Mufwene, Salikoko S. 2004a. Language birth and death. *Annual Review of Anthropology* 33: 201–22.
- Mufwene, Salikoko S. 2004b. Multilingualism in linguistic history: Creolization and indigenization. In Tej K. Bhatia and William C. Ritchie (eds.), *Handbook of Bilingualism*. Malden, MA: Blackwell, 460–88.
- Mufwene, Salikoko S. 2005. Language evolution: The population genetics way. In Günter Hauska (ed.), *Gene, Sprachen, und ihre Evolution*. Universitätsverlag Regensburg, 30–52.
- Mufwene, Salikoko S. 2007. Population movements and contacts in language evolution. *Journal of Language Contact – THEMA* 1: 63–91. [www.jlc-journal.org](http://www.jlc-journal.org).
- Mufwene, Salikoko S. 2008. *Language Evolution: Contact, Competition and Change*. New York: Continuum.
- Muysken, Pieter 1988. Are Creoles a special type of language? In Frederick J. Newmeyer (ed.), *Linguistics: The Cambridge Survey*, vol. II. *Linguistic Theory: Extensions and Implications*. Cambridge University Press, 285–301.
- Muysken, Pieter 2000. *Bilingual Speech: A Typology of Code-Mixing*. Cambridge University Press.
- Muysken, Pieter and Norval Smith (eds.) 1986. *Substrata versus Universals in Creole Genesis*. Amsterdam: John Benjamins.
- Muysken, Pieter and Tonjes Veenstra 1994. Universalist approaches. In Arends, Muysken and Smith (eds.), 121–34.
- Mühlhäusler, Peter 1986. *Pidgin and Creole Linguistics*. Oxford: Blackwell.
- Mühlhäusler, Peter and Philip Baker 1996. English-derived contact languages in the Pacific in the 20th century. In Stephen A. Wurm, Peter Mühlhäusler and Darrel T. Tryon (eds.), *Atlas of Languages of Intercultural Communication in the Pacific, Asia and the Americas*. Berlin: Mouton de Gruyter, 497–522.
- Naro, Anthony J. 1978. A study on the origins of pidginization. *Language* 54.2: 314–49.
- Nettle, Daniel 1999. Is the rate of linguistic change constant? *Lingua* 108: 119–36.
- Newitt, Malyn 2005. *A History of Portuguese Overseas Expansion 1400–1668*. London/ New York: Routledge.
- Newmeyer, Frederick J. 2000. *Language Form and Language Function*. Cambridge MA: MIT Press.
- Newmeyer, Frederick J. 2005. *Possible and Probable Languages: A Generative Perspective on Linguistic Typology*. Cambridge University Press.
- Nicol, John 1822. *The Life and Adventures of John Nicol*. Edinburgh: William Blackwood.
- Niebor, Herman J. 1919. *Slavery as an Industrial System: Ethnological Researches*. 2nd edn. The Hague: Nijhoff.
- Noble, Charles F. 1762. *A Voyage to the East Indies in 1747 and 1748*. London: Becket and Dehondt.



- Nordhoff, Sebastian 2009. A grammar of Upcountry Sri Lanka Malay. Unpublished PhD thesis, University of Amsterdam.
- Norman, Jerry 1988. *Chinese*. Cambridge University Press.
- Nothofer, Bernd 1995. The history of Jakarta Malay. *Oceanic Linguistics* 34.1: 87–97.
- Ochs, Elinor 1986. From feelings to grammar: A Samoan case study. In Bambi Schieffelin and Elinor Ochs (eds.), *Language Socialization across Cultures*. Studies in the Cultural Foundations of Language. Cambridge University Press, 251–72.
- Ochs, Elinor 1996. Linguistic resources for socializing humanity. In John Gumperz and Stephen Levinson (eds.), *Rethinking Linguistic Relativity*. Cambridge University Press, 407–38.
- Ochs, Elinor and Bambi Schieffelin 1984. Language acquisition and socialization: Three developmental stories and their implications. In Richard A. Shweder and Robert A. Levine (eds.), *Culture Theory: Essays in Mind, Self and Emotion*. New York: Cambridge University Press, 276–320.
- Ochs, Elinor and Bambi Schieffelin 2001. Language acquisition and socialization: Three developmental stories and their implications. In Duranti (ed.), 263–301.
- Olson, David R. 1994. *The World on Paper: The Conceptual and Cognitive Implications of Writing and Reading*. Cambridge University Press.
- Omoniyi, Tope and Goodith White (eds.) 2006. *The Sociolinguistics of Identity*. New York: Continuum.
- Ostler, Nicholas 2005. *Empires of the World: A Language History of the World*. New York: Harper Collins.
- Oyama, Susan 1985. *The Ontogeny of Information: Developmental Systems and Evolution*. Cambridge University Press.
- Paauw, Scott 2003. What is Bazaar Malay? The 5th International Symposium on Malay/Indonesian Linguistics, Nijmegen, June 2003.
- Pagel, Mark, Quentin D. Atkinson and Andrew Meade 2007. Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature* 449.11: 717–20.
- Pakir, Anne 1986. A linguistic investigation of Baba Malay. Unpublished PhD thesis, University of Hawaii.
- Pakir, Anne 1991. The range and depth of English-knowing bilinguals in Singapore. *World Englishes* 10.2: 167–79.
- Pan, Lynn (ed.) 2000. *The Encyclopedia of the Chinese Overseas*. Singapore: Chinese Heritage Centre/Cambridge, MA: Harvard University Press.
- Panikkar, Kavalam M. 1953. *India and the Indian Ocean: An Essay on the Influence of Sea Power on Indian History*. London: Allen & Unwin.
- Parkin, David 1994. *Continuity and Autonomy in Swahili Communities: Inland Influences and Strategies of Self-Determination*. London: SOAS.
- Parmentier, Raoul 1883. *Le Discours de la Navigation de Jean et Raoul Parmentier de Dieppe*. Paris: Ernest Leroux.
- Paviot, Jacques 2005. Trade between Portugal and the Southern Netherlands in the 16th century. In Ernst van Veen and Leonard Blussé (eds.), *Rivalry and Conflict: European Traders and Asian Trading Networks in the 16th and 17th Centuries*. Leiden: CNWS publications, 24–34.
- Pavlenko, Aneta and Adrian Blackledge 2004. Introduction: New theoretical approaches to the study of negotiation of identities in multilingual contexts. In Aneta

- Pavlenko and Adrian Blackledge (eds.), *Negotiation of Identities in Multilingual Contexts*. Bilingual Education and Bilingualism. Clevedon: Multilingual Matters, 1–33.
- Pearson, Michael 2003. *The Indian Ocean*. London: Routledge.
- Pierrehumbert, Janet 2001. Exemplar dynamics: Word frequency, lenition and contrast. In Bybee and Hopper (eds.), 137–58.
- Pierrehumbert, Janet 2002. Word-specific phonetics. In Carlos Gussenhoven and Natasha Warner (eds.), *Laboratory Phonology 7*. Berlin/New York: Mouton de Gruyter, 101–40.
- Pigafetta, Antonio 1524. *First Voyage Round the World*. Transl. by J.A. Robertson 1969. Manila: Filipiniana Book Guild.
- Pires, Tomé 1515 [reprint 1944]. *Suma Oriental*. Transl. by A. Cortesão. London: Hakluyt Society.
- Platt, John 1975. The Singapore English speech continuum and its basilect ‘Singlish’ as a ‘Creoloid’. *Anthropological Linguistics* 17.7: 363–74.
- Platt, John and Heidi Weber 1980. *English in Singapore and Malaysia*. Kuala Lumpur: Oxford University Press.
- Poyen-Bellisle, Rene 1894. *Le Sons et les Formes du Creoles des Antilles*. Unpublished PhD thesis, University of Chicago.
- Pratt, Mary Louise 1987. Linguistic utopias. In Nigel Fabb, Derek Attridge, Alan Durant, and Colin Maccabe (eds.), *The Linguistics of Writing*. Manchester University Press, 48–66.
- Ptak, Roderick 1998. *China and the Asian Sea: Trade, Travel and Visions of the Other (1400–1750)*. Aldershot: Variorum.
- Quennell, Peter (ed.) 1975. *Memoirs of William Hickey*. London: Routledge & Kegan Paul.
- Rabibhadana, Akin 1969. *The Organization of Thai Society in the Early Bangkok Period, 1792–1873*. Ithaca: Cornell University Southeast Asia Program.
- Rafferty, Ellen 1987. Word order in intransitive clauses in High and Low Malay of the late Nineteenth Century. In Russell S. Tomlin (ed.), *Coherence and Grounding in Discourse: Typological Studies in Language*. Amsterdam: John Benjamins, 361–82.
- Raffles, Thomas Stamford 1817. *The History of Java*, 2 vols. London: Back, Parbury & Allen and John Murray. Reprinted 1965, 1978 in Kuala Lumpur: Oxford University Press.
- Rampton, Ben 1990. Displacing the native speaker: Expertise, affiliation and inheritance. In Roxy Harrison and Ben Rampton (eds.), *The Language and Ethnicity Reader*. London: Routledge, 107–12.
- Rampton, Ben 2001. Creole metaphors in cultural analysis: On the limits and possibilities of (socio)linguistics. *Working Papers in Urban Language & Literacies*, no. 16. King’s College London.
- Raven-Hart, Rowland (ed.) 1953. *Germans in Dutch Ceylon, by Von der Behr (1668), Herport (1669), Schweitzer (1682) and Fryke (1692)*. Colombo National Museum and Ceylon Government Press.
- Ray, Haraprasad 1995. An enquiry into the presence of the Chinese in South and South-East Asia after the voyages of Zheng He in early fifteenth century. In Kuzhippalli S. Mathew (ed.), *Mariners, Merchants and Oceans: Studies in Maritime History*. New Dehli: Manohar, 97–110.

- Reid, Anthony 1980. The structure of cities in Southeast Asia, 15th to 17th centuries. *Journal of Southeast Asian Studies* 11.2: 235–50.
- Reid, Anthony (ed.) 1983. *Slavery, Bondage and Dependency in Southeast Asia*. New York: St. Martin's Press.
- Reid, Anthony 1988. *Southeast Asia in the Age of Commerce 1450–1680: The Lands Below the Wind*, vol. I. New Haven: Yale University Press.
- Reid, Anthony 1993a. *Southeast Asia in the Age of Commerce 1450–1680: Expansion and Crisis*, vol. II. New Haven: Yale University Press.
- Reid, Anthony (ed.) 1993b. *Southeast Asia in the Early Modern Era: Trade, Power and Belief*. Ithaca: Cornell University Press.
- Reid, Anthony 1996. Flows and seepages in the long-term Chinese interaction with Southeast Asia. In Anthony Reid (ed.), *Sojourners and Settlers: Histories of Southeast Asia and the Chinese*. Asian Studies Association of Australia: Allen & Unwin, 15–50.
- Reid, Anthony 2000. *Charting the Shape of Early Modern Southeast Asia*. Singapore: Institute of Southeast Asian Studies.
- Reimers, Edmund (ed.) 1932. *Memoirs of Ryckloff van Goens governor of Ceylon delivered to his successors Jacob Hustaart on December 26, 1663 and Ryckloff van Goens the Younger on April 12, 1675*. Colombo: Ceylon Government Press.
- Reinecke, John E. 1937. Marginal languages: A sociological survey of the creole languages and trade jargons. Unpublished PhD thesis, Yale University.
- Reinecke, John E., Stanley M. Tsuzaki, David de Camp, Ian F. Hancock and R. Wood 1975. *A Bibliography of Pidgin and Creole Languages*. Honolulu: The University Press of Hawaii.
- Renfrew, Colin 1987. *Archaeology and Language: The Puzzle of the Indo-European Origins*. London: Jonathan Cape.
- Rickford, John 1985. Ethnicity as a sociolinguistic boundary. *American Speech* 60: 99–125.
- Ricklefs, Merle C. 1974. *Jogjakarta under Sultan Mangkubumi, 1749–1779: A History of the Division of Java*. London/New York: Oxford University Press.
- Ricklefs, Merle C. 1991. *A History of Modern Indonesia since c. 1300*. 2nd edn. London: Macmillan.
- Roberts, Sarah J. 1998. The role of diffusion in the genesis of Hawaiian Creole. *Language* 74: 1–39.
- Roberts, Sarah J. 1999. The TMA system of Hawaiian Creole and diffusion. In John Rickford and Suzanne Romaine (eds.), *Creole Genesis, Attitudes and Discourse: Studies Celebrating Charlene J. Sato*. Amsterdam/Philadelphia: John Benjamins, 45–70.
- Roberts, Sarah J. 2000. Nativization and genesis of Hawaiian Creole. In John McWhorter (ed.), *Language Change and Language Contact in Pidgins and Creoles*. Amsterdam/Philadelphia: John Benjamins, 257–300.
- Roe, Michael (ed.) 1967. *The Journals and Letters of Captain Charles Bishop on the North-West Coast of America, in the Pacific and in New South Wales, 1794–1799*. Cambridge University Press.
- Rosa, Walter 2005. Portuguese land ordinance and urbanizing strategies for Asia. In Ernst van Veen and Leonard Blussé (eds.), *Rivalry and Conflict: European Traders and Asian Trading Networks in the 16th and 17th Centuries*. Leiden: CNWS publications, 55–70.

- Ross, Malcolm 1996. Contact-induced change and the comparative method: Cases from Papua New Guinea. In Mark Durie and Malcolm Ross (eds.), *The Comparative Method Revisited: Regularity and Irregularity in Language Change*. New York: Oxford University Press, 180–217.
- Rudolph, Jürgen 1998. *Reconstructing Identities: A Social History of the Babas in Singapore*. Aldershot/Vermont: Ashgate Publishing Co.
- Sachs, Jeffrey 2005. *The End of Poverty*. New York: Penguin.
- Sahlin, Marshall 1981. *Historical Metaphors and Mythical Realities: Structure in the Early History of the Sandwich Islands Kingdom*. Ann Arbor: University of Michigan Press.
- Saldin, B. D. K. 2003. *Portrait of a Sri Lankan Malay*. Kurunegala: Nihon Printers.
- Sampson, Geoffrey 1997. *Educating Eve: The Language Instinct Debate*. London: Cassel.
- Sardesai, D. R. 1997. *Southeast Asia: Past and Present*. 3rd edn. Chiang Mai: Silkworm Books.
- Sasi Rekha d/o Muthiah 2007. A description of Singapore Indian Malay. Unpublished MA thesis, National University of Singapore.
- Sastri, Nilakanta K. A. 1976. *A History of South India: From Prehistoric Times to the Fall of Vijaynagar*. 4th edn. Madras: Oxford University Press.
- Schieffelin, Bambi 1990. *The Give and Take of Everyday Life: Language Socialization of Kaluli Children*. Cambridge University Press.
- Schieffelin, Bambi and Rachel C. Doucet 1994. The 'real' Haitian Creole: Ideology, metalinguistics and orthographic choice. *American Ethnologist* 21.1: 177–201.
- Schiffman, Harold 1999. *A Reference Grammar of Spoken Tamil*. Cambridge University Press.
- Schleicher, August (1861–1862). *Compendium der Vergleichenden Grammatik der Indogermanischen Sprachen*, 2 vols. Weimar: Hermann Böhlau.
- Schleicher, August 1863. *Die Darwinsche Theorie und die Sprachwissenschaft*. Weimar: H. Böhlau.
- Schmidt, Johannes 1872. *Die Verwandtschaftsverhältnisse der Indogermanischen Sprachen*. Weimar: H. Böhlau.
- Schneider, Edgar 2003. The dynamics of New Englishes: From identity construction to dialect birth. *Language* 79.2: 233–81.
- Schuchardt, Hugo 1883. Kreolische Studien III. Über das Indoportugiesische von Diu. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaft zu Wien*. 103: 3–18.
- Schumann, John H. 1978a. The acculturation model for second language acquisition. In Rosario Gingras (ed.), *Second Language Acquisition and Foreign Language Teaching*. Washington DC: Center for Applied Linguistics.
- Schumann, John H. 1978b. *The Pidginization Process: A Model for Second Language Acquisition*. Rowley, MA: Newbury House.
- Schweitzer, Christopher 1931. *Reise nach Java und Ceylon 1675–1682*. Haag: Nijhoff.
- Searle, John 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press.
- Selbach, Rachel 2007. Lingua Franca of the Mediterranean (1350–1830): Finding Nessie? In Sture Ureland, Anthony Lodge and Stefan Pugh (eds.), *Language Contact and Minority Languages on the Littorals of Europe*, vol. V. Studies in Eurolinguistics. Berlin: Logos, 149–60.
- Selby, Anne and Stephen Selby 1997. China Coast Pidgin English. *Journal of the Hong Kong Branch of the Royal Asiatic Society* 35: 113–41.

- Shellabear, Rev. William G. 1913. Baba Malay: An introduction to the language of the Straits-born Chinese. *Journal of the Straits Branch of the Royal Asiatic Society* **65**: 49–63.
- Shi, Dingxu 1986. Chinese Pidgin English: Its origins and linguistic features. Ms. Department of Linguistics, University of Pittsburgh.
- Shi, Dingxu 1991. Pidgin English: Its origin and linguistic features. *Journal of Chinese Linguistics* **19**: 343–7.
- Siegel, Jeff 1997. Mixing, levelling, and pidgin/creole development. In Arthur K. Spears and Donald Winford (eds.), *The Structure and Status of Pidgins and Creoles*. Creole Language Library 19. Amsterdam/Philadelphia: John Benjamins, 111–50.
- Siegel, Jeff 2000. Substrate influence in Hawai'i Creole English. *Language in Society* **29**: 197–236.
- Siegel, Jeff 2004. Morphological elaboration. *Journal of Pidgin and Creole Languages* **19.2**: 333–62.
- Siegel, Jeff 2007. Recent evidence against the Language Bioprogram Hypothesis: The pivotal case of Hawai'i Creole. *Studies in Language* **31.1**: 51–88.
- Silva, A. W. L. 2003. *Teach yourself Tamil: A Complete Course for Beginners*. Kandana, Sri Lanka: Pubudu Printers.
- Silva, Renelde J. B. da 2001. *A Identidade Macaense: The Macanese Identity*. Macau: Instituto Internacional de Macau.
- Silverstein, Michael 1976. Hierarchy of features and ergativity. In R. M. W. Dixon (ed.), *Grammatical Categories in Australian Languages*. Canberra: Australian Institute of Aboriginal Studies, 112–71.
- Silverstein, Michael 1979. Language structure and linguistic ideology. In Paul R. Clynes, William F. Hanks and Carol L. Hofbauer (eds.), *The Elements: A Parasession on Linguistic Units and Levels*. Chicago Linguistics Society, 193–247.
- Singler, John Victor 1984. The homogeneity of the substrate as a factor in pidgin/creole genesis. *Language* **64.1**: 27–51.
- Singler, John Victor 1996. Theories of creole genesis, sociohistorical considerations, and the evaluation of evidence: The case of Haitian Creole and the relexification hypothesis. *Journal of Pidgin and Creole Languages* **11.2**: 185–230.
- Skinner, William G. 1960. Change and persistence in Chinese culture overseas: A comparison of Thailand and Java. *Journal of the South Seas Society* **16.1/2**: 86–100.
- Skinner, William G. 1996. Creolized Chinese societies in Southeast Asia. In Anthony Reid (ed.), *Sojourners and Settlers: Histories of Southeast Asia and the Chinese*. Asian Studies Association of Australia: Allen & Unwin, 51–93.
- Skuttnab-Kangas, Tove 2000. *Linguistic Genocide in Education – or Worldwide Diversity in Human Rights?* Mahwah, NJ: Lawrence Erlbaum.
- Slomanson, Peter 2006. Sri Lankan Malay morphosyntax: Lankan or Malay? In Ana Deumert and Stephanie Durrleman (eds.), *Structure and Variation in Language Contact*. Amsterdam/Philadelphia: John Benjamins, 135–58.
- Smith, Geoff P. 1995. Language choice in an acutely multilingual society: Communication and development in Papua New Guinea. International Conference on Language in Development: The Stakeholders' Perspectives. Denpasar, Bali, 10–12 April 1995.
- Smith, Geoff P. 2004. Chinese language sources for Chinese Pidgin English: What we know and what we need to know. *Hong Kong Journal of Applied Linguistics* **9.2**: 72–9.
- Smith, Geoff P. to appear. CPE pronouns revisited. *Hong Kong Journal of Applied Linguistics*.

- Smith, Geoff P. and Stephen Matthews (eds.) 2005. Special Issue on *Chinese Pidgin English: Texts and Contexts*. *Hong Kong Journal of Applied Linguistics* 10.1.
- Smith, Ian 2003. The provenance and timing of substrate influence in Sri Lanka Malay: Definiteness, animacy and number in accusative case marking. The South Asian Language Analysis Roundtable XXIII, Austin, Texas, 10–12 October 2003.
- Smith, Ian and Scott Paauw 2006. Sri Lanka Malay: Creole or convert? In Ana Deumert and Stephanie Durrleman (eds.), *Structure and Variation in Language Contact*. Amsterdam/Philadelphia: John Benjamins, 159–81.
- Smith, Ian, Scott Paauw and B.A. Hussainmiya 2004. Sri Lanka Malay: The state of the art. In R. Singh (ed.), *Yearbook of South Asian Languages 2004*. Berlin/New York: Mouton de Gruyter, 197–215.
- Sneddon, James 1996. *Indonesian Reference Grammar*. St. Leonards, NSW: Allen & Unwin.
- Soothill, William 1925. *China and the West: A Sketch of their Intercourse*. Oxford University Press.
- Sourjah, M.A. 2005. Malays in Sri Lanka came with their women-folk. *Sri Lanka Virtual Library*. [www.lankalibrary.com/cul/malays.htm](http://www.lankalibrary.com/cul/malays.htm).
- Souza, George B. 2005. The Portuguese merchant fleet at Macau in the seventeenth and eighteenth centuries. In Ernst van Veen and Leonard Blussé (eds.), *Rivalry and Conflict: European Traders and Asian Trading Networks in the 16th and 17th Centuries*. Leiden: CNWS publications, 342–69.
- Spolsky, Bernard 1989. *Conditions for Second Language Learning: Introduction to a General Theory*. Oxford University Press.
- Stavorinus, John S. 1798. *Voyages to the East Indies, by the Late John Splinter Stavorinus*. Transl. by S.H. Wilcocke. London: G. G. and J. Robinson.
- Stroud, Christopher 2000. Language and democracy: The notion of linguistic citizenship and mother tongue programmes. In Karsten Legère and Sandra Fitchat (eds.), *Talking Freedom: Language and Democratization in the SADC Region*. Windhoek: Macmillan, 67–74.
- Sutherland, Heather 1983. Slavery and the slave trade in South Sulawesi, 1660s–1800. In Reid (ed.), 263–85.
- Tamburello, Adolfo (ed.) 1983. La cultura occidentale nel Giappone Tokugawa, la mediazione olandese a russa nel 1800. *Il Giappone* 21: 5–23.
- Tan, Chee-Beng 1988a. *The Baba of Malacca. Culture and Identity of a Peranakan Community in Malaysia*. Petaling Jaya/Selangor: Pelanduk Publications.
- Tan, Chee-Beng 1988b. Structure and change: Cultural identity of the Baba of Melaka. *Bijdragen tot de Taal, Land- en Volkenkunde* 144: 297–314.
- Taylor, Jean G. 1983. *The Social World of Batavia: European and Eurasian in Dutch Asia*. Madison: The University of Wisconsin Press.
- Teixeira, Manuel 1994. The origin of the Macanese. *Review of Culture* 20: 157–61.
- Temple, Sir Richard C. (ed.) 1919. *The Travels of Peter Mundy, in Europe and Asia, 1608–1667*, vol. III, part I. Cambridge: Hakluyt Society.
- Than Tun (ed.) 1983. *The Royal Order of Burma, AD 1593–1885*. Kyoto University Center for Southeast Asian Studies.
- Thomason, Sarah and Terence Kaufman 1988. *Language Contact, Creolization and Genetic Linguistics*. Berkeley: University of California Press.
- Thomaz, Luis Felipe Ferreira Reis 1993. The Malay sultanate of Melaka. In Reid (ed.), 69–90.



- Thompson, Mary S. 1994. Vehicles all the way down. *Behavioral and Brain Sciences* **17**: 638.
- Thurgood, Elizabetha 2001. Language contact as reflected in Baba Malay forms of address and reference. In Graham W. Thurgood (ed.), *Papers from the 9th Annual Meeting of the Southeast Asian Linguistic Society 1999*. Arizona State University Monograph Series, 27–50.
- Thurgood, Elizabetha and Graham Thurgood 1996. Aspect, tense or aktionsart? The particle *ja* in Kristang (Melacca Creole Portuguese). *Journal of Pidgin and Creole Languages* **11.1**: 45–70.
- Tomás, Isabel 1988. Makista Creole. *Review of Culture* **15**: 33–46.
- Tomás, Isabel 1990. The life and death of a creole. *Review of Culture* **9**: 55–66.
- Tomasello, Michael 2003. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA/London, England: Harvard University Press.
- Tomasello, Michael, Anne Cale Kruger and Hilary Horn Ratner 1993. Cultural learning. *Behavioral and Brain Sciences* **16**: 495–552.
- Trask, Larry R. 1996. *Historical Linguistics*. London/New York: St. Martin's Press.
- Trudgill, Peter 2002. Linguistic and social typology. In J. K. Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.), *The Handbook of Language Variation and Change*. Malden MA/Oxford: Blackwell, 707–28.
- Tryon, Darrel T., Peter Mühlhäusler, and Philip Baker 1996. English-derived languages in the Pacific in the 19th century. In Stephen A. Wurm, Peter Mühlhäusler and Darrel T. Tryon (eds.), *Atlas of Languages of Intercultural Communication in the Pacific, Asia and the Americas*. Berlin: Mouton de Gruyter, 471–95.
- Tuuk, Herman N. van der 1856. Iets over de Hoog-Maleische Bijbelvertaling. *Bijdragen tot de Taal, Land- en Volkenkunde van Nederlandsch-Indië* **5**: 171–83.
- Valentyn, Francois 1726. *Oud en Nieuw Oost Indiën: Omstandig Verhaal van de Geschiedenissen en Zaaken het Kerkelyke ofte den Godsdienst Betreffende, zoo in... Amboina* Dordrecht: Johannes van Braam.
- Van Dyke, Paul 2002. Port Canton and the Pearl River Delta, 1690–1845. Unpublished PhD thesis, University of Southern California.
- Van Dyke, Paul 2005. *The Canton Trade: Life and Enterprise on the China Coast, 1700–1845*. Hong Kong University Press.
- Wang, Gungwu 1996. Sojourning: The Chinese experience in Southeast Asia. In Anthony Reid (ed.), *Sojourners and Settlers: Histories of Southeast Asia and the Chinese*. Asian Studies Association of Australia: Allen & Unwin, 1–14.
- Warren, James F. 1981. *The Sulu Zone 1768–1898. The Dynamics of External Trade, Slavery, and Ethnicity in the Transformation of a Southeast Asian Maritime State*. Singapore University Press.
- Weinreich, Uriel 1953. *Languages in Contact: Findings and Problems*. New York: Linguistics Circle of New York.
- Wenger, Etienne 1998. *Communities of Practice: Learning, Meaning and Identity*. Cambridge University Press.
- Whinnom, Keith 1971. Linguistic hybridisation and the special case of pidgins and Creoles. In Hymes (ed.), 91–116.
- Whorf, Benjamin Lee 1956. A linguistic consideration of thinking in primitive communities. In John B. Carroll (ed.), *Language, Thought and Reality: Selected Writings of Benjamin Lee Whorf*. Cambridge: MIT Press, 65–86.

- Williams, Samuel W. 1836. Jargon spoken at Canton. *Chinese Repository* 4: 428–35.
- Winford, Donald 2003. *An Introduction to Contact Linguistics*. Oxford: Blackwell.
- Winford, Donald 2005. Contact-induced changes: Classification and processes. *Diachronica* 22.2: 373–427.
- Wolff, Ekkard 2000. Language and society. In Bernd Heine and Derek Nurse (eds.), *African Languages: An Introduction*. Cambridge University Press, 298–347.
- Woolard, Kathryn A. 1999. Simultaneity and bivalency as strategies in bilingualism. *Journal of Linguistic Anthropology* 8.1: 3–29.
- Yap, Foong Ha and Shoichi Iwasaki 2003. From causative to passive: A passage in some East and Southeast Asian languages. In Eugene Casad and Gary Palmer (eds.), *Cognitive Linguistics and Non-Indo-European Languages*. Berlin: Mouton de Gruyter, 419–46.
- Yip, Virginia, and Stephen Matthews 2007. *The Bilingual Child: Early Development and Language Contact*. Cambridge University Press.

## HISTORICAL RESOURCES

- Ming Shi Lu*, transl. by Geoff Wade, Southeast Asia in the Ming Shi-lu: an open access resource, Singapore: Asia Research Institute and the Singapore E-Press, National



# Index

---

- accommodation, 41, 89, 103, 143, 150, 161, 179, 220, 224
- acculturation, 66, 93, 104, 149
- Adelaar, 8, 23, 54–59, 60–63, 69, 80, 129, 135, 142, 146, 162–164, 166, 171, 174, 177, 182–183, 222
- admixture, 3, 5, 8, 15, 19, 43, 61, 66, 79, 86, 89, 94, 97, 119–120, 128, 138–142, 150, 155, 158–161, 164, 226
- adstrate, 6, 92, 117, 130, 134, 137, 142, 145, 175, 178, 224
- agency, 4–6, 95, 114, 181, 194, 208, 217, 221, 224
- Ambon, 64, 124, 146
- Arends, 14, 213, 215–216
- Asian-Portuguese, 53, 74–75, 77–78, 148, 150, 153–154, 159, 161–162, 164, 184, 188, 191, 194, 206, 209, 217–218
- Baba Malay, 3, 8–9, 30, 37, 53, 60–66, 71–72, 122, 133, 136, 153, 158–159, 161–163, 165–167, 171–175, 177–182, 201, 205, 215–216, 222, 224
- causatives, 174
- demonstratives, 170
- ecology, 162
- evolution, 161–162, 178–179
- existential constructions, 171–172
- grammar, 166
- kena*, 175–176
- lexicon, 163–165, 179
- loanwords, 163
- morphology, 176–177
- negation, 172
- phonology, 163, 179
- pronouns, 165
- reduplication, 167–168
- substrate reinforcement, 176
- tones, 162–163
- topic, 166–167
- zero-copula, 171
- Batavia, 21, 34, 46, 67–68, 125, 182
- Bazaar Malay, 3, 8, 23, 35, 53–55, 59–64, 68–74, 78, 80, 89, 113, 125, 133, 136, 146, 150, 161–162, 166–168, 171–172, 176, 178–179, 182, 201–202, 205, 222, 227
- causatives, 174
- demonstratives, 170
- existential constructions, 172
- kena*, 175–176
- Melaka, 163
- morphology, 177
- phonology, 163
- possessive *punya*, 172–174
- pronouns, 165
- reduplication, 167–168
- relexification, 178
- Singapore, 64, 69–72, 162–163, 172–174, 179
- substrate reinforcement, 176
- topic, 167
- variation, 163, 168
- zero-copula, 171
- Bickerton, 71, 92
- bondage, 37, 39, 45, 51
- British East India Company, 42, 48, 66, 68–69, 196
- Bybee, 10, 105, 114–116
- Canton, 28, 77, 154, 159, 184, 187–198, 207–208, 210, 217
- Cantonese, 2, 6–7, 27–29, 50, 52, 69–70, 75, 78, 113, 153–154, 159, 164, 170, 182, 189–190, 194, 197–207, 211, 217
- Cape Malay, 64
- Chaudenson, 41, 72, 93, 210, 215, 217, 220, 224
- China Coast Pidgin, 6, 120, 184, 186, 188–189, 192–199, 201, 204–212, 216–219, 225
- and Makista, 205–206
- calques, 201
- Chinese Pidgin English, 184
- compounding, 200

- China Coast Pidgin (*Cont.*)  
 copula, 200–201  
 data sources, 197–198  
 ecology, 208  
 evolution, 208–209  
 existential constructions, 201  
 grammar, 197  
 in Macau, 207  
 lexicon, 198  
 locatives, 204  
 negation, 202  
 phonology, 198–199  
 possessive, 201  
 pronouns, 199  
 quantifiers, 204  
 reduplication, 205  
 serialization, 204  
 substrate influence, 197–198, 205  
 topic, 205  
 variation, 197  
 VP, 202
- Chitty Malay, 23, 61, 63
- Chola empire, 78
- Chomsky, 9, 17, 84, 101, 103, 117, 120
- city-ports, 6, 19, 21, 25, 27–30, 33–38,  
 40–41, 43, 48, 51–52, 54, 56, 65, 79,  
 184, 208
- Cocos Islands, 46, 64, 168
- Cocos Malay, 3, 16, 53, 63, 66, 68–69, 71–73,  
 122, 136, 162–166, 168, 172, 174–179,  
 182–183, 222  
 causatives, 175  
 demonstratives, 170  
 lexicon, 164  
 morphology, 177  
 passive, 175–177  
 phonology, 163  
 possessives, 173–174  
 pronouns, 164, 165  
 reduplication, 167  
 relative marker, 174  
 zero-copula, 171
- Collins, 28, 53–57, 59–60, 62–63, 163, 167
- colonization, 6, 43, 123, 143
- competition, 11, 34, 47, 100–101, 111, 137,  
 154, 184
- complexity, 86, 118, 142, 180, 195, 208, 223  
 complexification, 216, 227
- compradors, 191–192, 194, 196, 208
- concubinage, 36, 77, 153
- congruence, 5, 6, 12–13, 15, 72, 87, 90, 101,  
 111–113, 117, 123, 128, 130, 132–134,  
 136–137, 141, 144, 150, 160, 162–164,  
 167–168, 171–172, 175–176, 178–180,  
 182, 216, 219
- contact language formation, 19, 35, 65, 72,  
 78–79, 83, 97, 99, 103–105, 107, 110–114,  
 117–123, 128, 137, 141–151, 178, 181,  
 185, 188, 209, 213–217, 219–220,  
 222–225
- contact languages, 14, 29, 55, 62, 73, 75, 78,  
 81, 90, 95, 97, 105, 107, 120, 129, 148,  
 150, 210, 213, 215–217, 225
- contact Malay, 21, 69, 122, 161, 171,  
 174–175, 177–178, 214, 216–217, 221
- convergence, 8, 85, 139, 160, 176–177,  
 218, 220
- creativity, 226
- Creole Prototype, 118, 142
- Creoles, 80, 87, 90–91, 97, 118, 200, 214  
 Asian, 169  
 Asian-Portuguese, 181, 209  
 Caribbean, 216  
 creolization, 42, 92, 138–139, 141–143,  
 218  
 French, 41–42  
 Indonesian, 61  
 Malayo-Portuguese, 75  
 Portuguese, 75, *see also* Indo-Portugese  
 Surinamese, 112
- Croft, 4, 9, 11, 15, 87–89, 96, 98, 105–106,  
 108–109, 111, 115, 117–118, 213, 222,  
 225–226
- Darwin, 67, 84, 92, 98
- DeGraff, 13–14, 84, 86, 90, 118, 141, 147,  
 222, 224
- diffusion, 85
- divergence, 6, 11, 108, 161, 181
- Dutch East India Company, 44–45
- ecology, 4–5, 8–13, 16, 18–19, 29, 52, 57, 61,  
 68–71, 73, 78, 80, 92, 95, 97–98, 100–101,  
 103, 105, 107–108, 110–113, 117–120,  
 123, 125, 128, 140, 145–146, 149–151,  
 153–154, 157–158, 160–163, 172, 177,  
 179, 181–182, 184–185, 188–189,  
 193–194, 196, 202, 205, 207, 210, 214,  
 216, 219, 222–224, 226  
 internal, 112  
 monolingual, 97  
 multilingual, 3, 142–143, 148–149, 151,  
 160, 196, 210, 223–225  
 multilingual ecology, 13  
 pidgin, 185, 210, 225
- endogamy, 35, 37, 151, 180, 215
- Estado da Índia*, 74
- evolution, 1, 3, 5–6, 8, 10, 12–15, 35,  
 37, 40, 42, 53, 55, 59, 65–66, 70–72,  
 74–75, 78–80, 85, 87, 90–92, 94–97,

- 103, 105–108, 110–111, 113–115, 117,  
119–123, 134–146, 149–150, 157, 159,  
161, 170, 172, 177, 179, 181, 185, 188,  
193–194, 197–198, 205, 209–210,  
213–217, 219–222, 224–226
- exogamy, 35, 37, 151
- Faracías, 6, 41, 93, 193–194, 215
- Feyerabend, 81, 227
- frequency, 11–12, 110–111, 113, 115–117,  
119–120, 122–123, 131–134, 136, 140,  
177, 180–181, 204, 216, 218–219, 223
- Fujian, 24–25, 28–29, 65, 158
- Givón, 9–10, 114–115, 117, 216
- Guangdong, 25, 28
- Hall, 47, 104, 194, 197, 201, 208
- Haspelmath, 85, 95, 113, 115, 117, 135
- Hokkien, 2, 7–8, 52, 61–62, 64, 66, 70,  
78, 80, 113, 124, 146, 155, 158,  
161–168, 172–180, 182, 194,  
201, 205
- Hong Kong, 2, 80, 160, 195–196, 204, 211
- Hull, 98, 105–108, 110–111, 118–120
- Hussainmiya, 123–125, 134, 138–139,  
146–147
- hybridity, 15, 87  
cultural, 24, 79, 151, 155, 220  
ecological, 215  
genetic, 97, 221  
structural, 8, 73, 131, 170, 180
- identity, 126, 144, 148–149, 151–154,  
157–158, 160–161, 180–181, 197, 207,  
210, 220  
acts of, 150, 220  
adaptation, 128  
construction, 96  
intra-communal, 207  
loss, 220  
new formation, 5, 8, 15, 71, 104, 153,  
215, 220  
realignment, 144, 150
- identity alignment, 104, 122–123, 148–151,  
157–161, 181, 215, 220–221, 224
- indenture, 8
- Indian-Portuguese, 75
- Indonesian, 68, 78, 177  
Riau, 125
- Indo-Portuguese, 73, 140, 156, 162, 164, 169,  
182, 206
- intermarriage, 6, 37, 43, 122, 128, 138, 140,  
145–146, 153, 209, 215
- interpreters, 54, 152, 189, 192, 194
- Jakarta Malay, 63, 146
- jargon, 62, 139–140, 188, 209
- Java, 18, 21–22, 24–25, 27–31, 33, 36, 42, 44,  
46, 50, 57, 61, 63–66, 68–69, 72, 78, 122,  
124, 152, 155, 158, 161–162, 227
- Java Malay, 61, 63–64, 72, 78
- Java Port Malay, 63
- Javanese, 52–53, 57, 69, 78, 125, 146,  
161–162, 164, 227
- Javanese Malay, 162
- Kirinda, 127–128, 157
- Lass, 89, 95, 106
- Latin, 54–55, 57, 84, 86, 89
- Lightfoot, 12, 95
- lingua franca, 52, 54–55, 60–62, 64, 70,  
73–75, 91, 94, 125, 133, 141, 144, 146,  
158, 161, 166, 189, 192, 194, 206–207,  
209, 219
- linguists, 191–192, 194, 208
- Macau, 3, 21, 43–44, 73–77, 80, 150,  
153–154, 156, 159–160, 163,  
181–182, 184, 187–192, 195–198,  
206–208, 211
- maintenance, 4, 10, 13–14, 16, 66, 73, 96,  
104, 111, 114, 143, 161, 165, 179, 181,  
224, 227
- Makista, 3, 9, 37, 44, 72, 74–75, 78, 80,  
150–151, 154, 159–161, 178, 180–182,  
188, 194, 205, 207–210, 215–216, 218,  
222, 224
- and China Coast Pidgin, 198, 205
- as lingua franca, 192
- evolution, 162
- lexicon, 164
- morphology, 168
- Sinitic influence, 163
- substrates, 169–170
- zero-copula, 171
- Malay, 34, 53–55, 57, 59, 63–70, 72–73, 75,  
78, 129, 131, 150, 157  
classical, 59–60  
classification, 55–57, 59  
colloquial, 54, 60, 72, 132  
contact, 55–57, 60, 62, 133, 136, 141,  
152, 158  
contact Malay, 53, 56, 61, 72  
creole, 123, 138  
evolution, 55, 57, 59, 62, 71  
Indonesian, 136  
literary, 23, 54, 57, 59, 62, 64–65, 72  
low/high, 55–56  
pidgin, 60–61, 63

Malay (*Cont.*)

- pidgin-derived, 57, 59, 61–62, 64, 68, 71–73, 80, 125, 128–130, 133–135, 137, 140, 144–145, 161, 163, 165–167, 172, 174, 179, 183, 201, 222
- restructured, 60, 129
- standard, 68, 70
- typology, 129, 131, 135–136
- Malay/Indonesian, 8, 29, 35, 45, 54, 56, 60–61, 63, 65–66, 79, 124, 129, 142, 152–153, 155, 158, 161–162, 164, 167–168, 170, 174–179, 182, 220, 222
- Malayo-Portuguese, 73–75, 160, 162, 168–169, 180
- Matras, 35, 97, 111, 128, 135, 215, 223
- Mediterranean, 9, 22, 48, 54, 62, 74
- Melaka, 2, 8, 18, 21, 23, 27–34, 41–44, 46, 49–50, 54, 56, 61, 63–66, 68–69, 72–78, 150, 152–156, 161–164, 182, 214
- merchants, 49, 65, *see also* traders
  - as middlemen, 152
  - Babas, 64, 158
  - Chinese, 6, 25, 79, 153, 161
  - Genoese/Venetian, 55
  - Java, 24
  - language use, 208
  - Portuguese, 77
  - sojourners, 25, 36
  - Western, 6, 187, 189, 192
- Ming Dynasty, 22, 25
- mixed languages, 35, 85, 87, 89, 95, 97, 118, 128
- Moluccan Malay, 136
- Moluku, 124
- monolingualism, 15, 86, 88–89, 97, 119, 143, 148, 228
- monsoon, 6, 8, 18–20, 22, 25, 28, 52–54, 62, 65, 73, 78, 112, 125, 144, 152, 187, 213
- Moors, 18, 124, 138–140, 147
- morphological elaboration, 216
- Mufwene, 3–5, 10–14, 61–62, 66, 72, 74, 83, 85, 87–92, 94–98, 105–106, 111–113, 118, 133, 138, 141, 149, 153, 159, 209–210, 213, 215, 220, 227
- multilingualism, 5–6, 8, 15, 35, 65, 72, 88–89, 92, 97, 102, 125, 127, 143, 146, 148–149, 151, 153–154, 180, 209, 220–223, 225
- Muslim, 21–22, 25, 27–28, 30–31, 43–44, 47–48, 51, 56, 64, 66, 68–69, 125, 138, 146, 151, 155, 192

Ochs, 87, 102–104, 120, 219

Old Malay, 23, 57, 222

Papia Kristang, 75

- pidgins, 62, 71–72, 74, 80, 92, 104, 119, 142, 150, 184–186, 197–198, 200, 205, 209–210, 214, 216–217, 219, 225, *see also* Malay, pidgin(-derived)
- English, 194
- in creole formation, 71, 72, 92, 113, 141, 225
- Melanesian, 75
- Pacific, 205
- Portuguese, 74, 206, 210
- Russenorsk, 184

pilots, 191, 194

pluralism, 151

propagation, 12, 85, 93, 109–112

Qing Dynasty, 188

- relexification, 142, 147, 178, 182, 206
- replication, 4, 11, 89, 92, 105, 107–111, 113–114, 116, 118–121, 123, 128, 133–134, 142, 144, 181, 219, 224
- identical, 110, 119
- innovative, 110, 117, 119
- restructuring, 3, 6, 16, 86, 90–91, 116, 119, 123, 134–135, 137, 142, 147, 161, 168, 173, 177, 199, 216, 219, 224
- Ross, 8, 85

Schieffelin, 87, 90, 96, 102–103, 219

- selection, 4, 11, 13, 92, 100–101, 105–113, 115, 118–120, 123, 133, 136–137, 142, 144, 147, 150, 167, 172, 175, 177, 179, 181, 213, 215–216, 218–219, 224

- simplicity, 179–180, 186, 219, 223
- simplification, 60, 71–72, 74, 104, 117, 180, 186, 215–218, 219, 224

Singapore, 1, 3, 18, 52–53, 59–61, 63–64, 66, 68–73, 89, 94, 110, 113, 150, 152, 155, 159, 163, 167, 182

Singlish (Singapore Colloquial English), 1, 3, 16, 52, 62, 70–71, 89, 94, 113, 117–118, 135, 168, 201

Sinhala, 13, 15–16, 113, 117, 122, 125, 127–144, 146–147, 150–151, 157, 216

Sinitic, 3, 6, 52–53, 56, 62–63, 65, 70–73, 75, 78, 80, 94, 101, 113, 117, 158–159, 161–164, 167–174, 177–180, 182, 184, 188, 194, 198–206, 209, 218, 222

Sino-Javanese, 22, 28–29, 34, 50, 79, 152

slavery, 6, 15, 34–35, 37–38, 40–41, 45, 51, 122, 214

socialization, 88, 101–104, 149, 224, 226

spice trade, 23, 25, 29, 43, 45, 186

- Sri Lanka Malay, 3, 13, 15–16, 62, 64, 73,  
 90–92, 98, 103–104, 112–113, 118,  
 122–151, 157–160, 177, 215–217, 220,  
 222–224  
 agents, 130  
 alignments, 131  
 animacy, 132  
 case, 111, 117, 129, 133–134, 216  
 definite object marker, 130  
 Diaspora, 125–127  
 ecology, 123, 128  
 evolution, 122–123, 128  
 experiencers, 130  
 Java, 127–128  
 lexicon, 129  
 morphology, 129  
 multilingualism, 127–128  
 NP, 137  
 tense/aspect, 134–137  
 typology, 122  
 VP, 136–137  
 word order, 13, 216  
 Sri Vijaya, 23, 30–31, 57, 59, 78  
 substrate, 98, 138, 140–141, 176, 182, 205,  
 217, 219, 225  
 Sumatra, 21, 23, 29, 31, 57, 65, 67, 155  
 Sundanese, 52–53, 78, 162  
 superstrate, 118, 141  
 Tamil, 13, 15–16, 53–54, 70, 78, 89, 113, 117,  
 122, 125, 127, 129, 130–144, 151, 157  
 Lankan, 128–129, 140, 150  
 Tamil bias, 138, 141  
 Tomasello, 10, 105, 114, 220  
 traders, 23, 29, 34, *see also* merchants  
 Chinese, 28, 36, 65, 70, 79, 161, 217  
 Hokkien, 61, 78  
 Javanese, 23  
 Melaka, 32  
 Muslim, 28, 29  
 North African, 55  
 Portuguese, 44, 75–76, 218  
 Siam, 31  
 Western, 44, 71, 73–74, 154, 187, 189,  
 194, 208  
 transmission, 6, 10, 15, 41, 85, 87, 89–90,  
 96, 99, 101, 105, 114, 119, 144, 149, 181,  
 215, 224  
 abrupt, 90  
 break in, 72  
 direct, 91, 221  
 horizontal, 85, 222  
 informal, 5, 14–15, 89, 105, 118, 226  
 linear, 144, 219, 222  
 monolingual, 99, 224  
 multilingual, 100–101, 119, 144, 224,  
 227  
 ‘normal’, 87–88, 90, 97  
 normative, 86  
 trilingualism, 6, 142, 144  
 trilingual, 122–123, 127–128, 137, 141,  
 143, 148, 158, 216, 226  
 typological matrix, 72, 112–113, 117–121,  
 123, 125, 128, 133–134, 136–138, 142,  
 145, 147, 160–162, 168, 171, 177–180,  
 182, 216–219, 222, 224  
 Zheng He, 18, 25, 27, 47, 50, 65